



The UNEP/ SETAC LCI Database Registry and the data Format Converter

International Open Workshop
“Worldwide efforts on LCA Databases“, 8 February 2010

Andreas Ciroth, Mark Goedkoop

Database Registry and Format Converter, February 8 2010

Outline

- intro: what are the registry and the format converter
- data format converter
- UNEP/SETAC database registry
- spirit of the registry and the format converter
- conclusion



Database Registry and Format Converter, February 8 2010

1. Introduction:

What are the Database Registry and the data
Format Converter (and what are they not)



Database Registry and Format Converter, February 8 2010

The openLCA data format converter

- open source software tool to convert between major LCA data formats
- first two releases 2007 (ELCD, EcoSpold 1, ISO1448/IMI formats)
- quite popular

openlca_converter	41.2 MB	2007-09-15	1,806	
openlca_converter_1.1	20.7 MB	2007-08-15	1,464	
converter1.0	20.6 MB	2007-05-04	342	

downloads from sourceforge site, www.sourceforge.net/projects/openlca,
per Jan 25, 2010

- supported by UNEP/SETAC life cycle initiative, PRé Consultants, and PE International
- designed and implemented by GreenDeltaTC



Database Registry and Format Converter, February 8 2010

The openLCA data format converter, 2.0

- update of the first version, also as open source software, released at the same location
- formats covered: ILCD, EcoSpold 1, EcoSpold 2
- also covered: nomenclature mapping
- collaboration with JRC (→ ILCD) and theecoinvent centre (→ EcoSpold 2) to „smoothen“ conversion problems
- supported so far by UNEP/SETAC life cycle initiative, ecoinvent centre
- designed and implemented by GreenDeltaTC



Database Registry and Format Converter, February 8 2010

The database registry

A web system to

- connect LCA data providers and users that search for LCA data, and
- become a "focal point" where information about international, sectoral, national, commercial and non commercial databases for life cycle approaches worldwide can be found.
- „open source“ part: availability to upload single data sets, by anyone (Mark Goedkoop)



2. The Data Format Converter



Data Format Converter



First Announcement:

Ciroth, A.: A new open source LCA software, 7th Ecobalance conference, Tsukuba, presentation and conference proceedings, pp. 427, Nov. 14 – Nov. 16, 2006



Format Converter: 3 key ideas

- 1 There are different LCA data formats that are relevant in practice (EcoSpold, ILCD, ...)
- 2 Users need to be able to „switch“ between them
 - data exchange, not all is available in one format
 - formats have different pros and cons in use, fit better for one or the other application
- 3 Open source (in common definition): making the conversion transparent, re-usable, extendable



Format Converter: Status

- Format Converter 2.0 to cover
 - * ILCD (replaced ELCD format summer 2009, by JRC)
 - * EcoSpold 1 (practical relevance)
 - * EcoSpold 2 (currently being finalised by ecoinvent centre)
- Format Converter 2.0 to provide also nomenclature mapping

elementary flow names	
ecoinvent	ELCD
Ethane, 1,1,2-trichloro-1,2,2-trifluoro-, CFC-113	CFC-113
Ethylene diamine	ethylenediamine
Gypsum, in ground	gypsum
PAH, polycyclic aromatic hydrocarbons	polycyclic aromatic hydrocarbons



Format Converter: Status

- converter 2.0: first release soon (some coordination work with JRC, ecoinvent)
- both ILCD and EcoSpold 02 are now recognising each other, and some major conversion issues could be solved
- EcoSpold02 more complex than EcoSpold 1
- nomenclature mapping requires maintenance, „a living structure“ (new flows, ...)
 - community
 - mapping of objects, not of fields and attributes
 - continuous improvement



Some conversion issues, 1: Text field lengths

EcoSpold:

In EcoSpold, the following data types for Strings (text) exist

TString20 (20 chars)
 TString30 (30 chars)
 TString40 (40 chars)
 TString80 (80 chars)
 TString120 (120 chars)
 TString255 (255 chars)

ILCD:

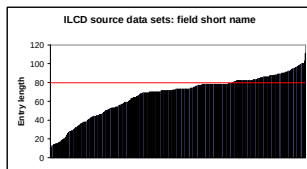
In ILCD, the following types exist for text:

500 (500 chars),
 1000 (1000 chars),
 free text (unlimited)



Some conversion issues, 1: Text field lengths

ILCD source data set, field short name (500) ->
EcoSpold source element, field short name (80)



Red line: EcoSpold field limit; black: ELCD database content fields

→ roughly 1/3 of ELCD field entries are longer than allowed in EcoSpold (for the source dataset name)

Some conversion issues, 2:
Elementary flow categories

ILCD			ecoinvent 2 update	
3 Emissions	Emissions to air	Emissions to lower stratosphere and upper troposphere	air(output to)	lower stratosphere + upper troposphere
5 Emissions	Emissions to air	Emissions to urban air close to ground	air(output to)	high population density
2 Emissions	Emissions to air	Emissions to air, unspecified (long term)	air(output to)	low population density/long term
4 Emissions	Emissions to air	Emissions to non-urban air or from high stacks	air(output to)	low population density
1 Emissions	Emissions to air	Emissions to air, unspecified	air(output to)	unspecified
13 Resources	Resources from air	Renewable element resources	natural resource(input of)	in air
14 Resources	Resources from air	Renewable energy resources	natural resource(input of)	in air
15 Resources	Resources from air	Renewable material resources	natural resource(input of)	in air
17 Resources	Resources from biosphere	Renewable energy resources	natural resource(input of)	biotic
18 Resources	Resources from ground	Non-renewable element resources	natural resource(input of)	in ground
19 Resources	Resources from ground	Non-renewable energy resources	natural resource(input of)	in ground
20 Resources	Resources from ground	Non-renewable material resources	natural resource(input of)	in ground
21 Resources	Resources from ground	Renewable energy resources	natural resource(input of)	in ground
22 Resources	Resources from ground	Renewable material resources	natural resource(input of)	in ground
23 Resources	Resources from water	Non-renewable element resources	natural resource(input of)	in water
24 Resources	Resources from water	Non-renewable energy resources	natural resource(input of)	in water
25 Resources	Resources from water	Non-renewable material resources	natural resource(input of)	in water
26 Resources	Resources from water	Renewable material resources	natural resource(input of)	in water
6 Emissions	Emissions to soil	Emissions to agricultural soil	soil(output to)	agricultural
8 Emissions	Emissions to soil	Emissions to soil, unspecified	soil(output to)	unspecified
7 Emissions	Emissions to soil	Emissions to non-agricultural soil	soil(output to)	forestry
10 Emissions	Emissions to water	Emissions to sea water	water(output to)	industrial
9 Emissions	Emissions to water	Emissions to fresh water	water(output to)	ground
8 Emissions	Emissions to water	Emissions to fresh water	water(output to)	ground, long term
11 Emissions	Emissions to water	Emissions to water, unspecified	water(output to)	unspecified
12 Emissions	Emissions to water	Emissions to water, unspecified (long term)	water(output to)	unspecified
16 Resources	Resources from biosphere		natural resource(input of)	land
			Economic(input of)	unspecified
			Social(input of)	unspecified

Some conversion issues, 2:
Elementary flow categories

...related to nomenclature and flow mapping!



Format Converter: path forward

- (try to) coordinate with ecoinvent & JRC conversion decisions
- release format converter 2.0 as open source
- establish a UNEP/SETAC project for maintaining the converter
- continuous improvement, adaptations of the converter
- versions integrated in websites and in other LCA software possible



3. UNEP/SETAC LCI database registry



UNEP/SETAC database registry

- aim:
 - * connect LCA data providers and users that search for LCA data, and
 - * become a "focal point" where information about international, sectoral, national, commercial and non commercial databases for life cycle approaches worldwide can be found.
- general support by a survey in 2008
- now a "living system" is implemented



Overview of the UNEP/SETAC database registry project

- addressees:

- * data providers
(no preconditions for providing data)
- * data users
(Life Cycle practitioners, scientists, ... , in search of suitable data).

- idea:

- * connect available data of high quality (and providers of these data)
- * and users that look for data, worldwide



UNEP/SETAC database registry project guiding principle:

Data quality understood as in ISO 14040.

→ ISO 14040 defines data quality as
“characteristics of data that relate to their ability to satisfy
stated requirements”

You could also say: “beauty lies in the eye of the beholder”.

!



UNEP/SETAC database registry: Consequence of the guiding principle

It is not an absolute property of a data set to be of
high or low quality.

Rather, data quality is result of properties of a
data set and of what users require from the data
set.



UNEP/SETAC database registry: Consequence of the guiding principle



Strange combinations?

Worth 1000 com, taken from <http://www.vinaminh.com/?p=1223>



UNEP/SETAC database registry: Consequence of the guiding principle

Or rather: Possibility to combine different data
sources, from different background.

Users are not „locked“ in one system.
Data Providers can reach many different users.



Data quality not an absolute property:

data quality is result of how well
stated requirements from users and
properties of data match

- responsibility to allow good ways to state requirements
- need to state requirements (data users) and to provide
data properties (data providers, project team?)
- responsibility to use the data as specified (users)
- need for control mechanisms?



UNEP/SETAC database registry project further principles

- fit for organic growth (scalable system instead of starting oversized)
 - openness (almost every Life Cycle data source)
 - flexibility
- data might be accessible

a) on a data source level
b) on a process group level
c) on a process meta data level
d) on a process level



UNEP/SETAC database registry current implementation

www.lca-data.org



Current implementation:

- web server, search engine, file upload implemented
- social collaboration website framework (→ facebook asf.)
- currently built:
 - * data stock (permission: ecoinvent meta data, ELCD complete, Gabi databases meta data, NREL complete, ...)
 - * open source data stock: some 100 complete ecoinvent data sets, Australian database, ...
 - * search logic (time, geography, technology, more)



What's next?

- Refine search logic
- Build organisational structure
 - * Board / council:
 - agreed members: ecoinvent, GaBi, Japanese Database, NREL, Worldsteel... (many thanks!)
 - call for additional members (overall 8-10 members)
 - * Editorial working groups (→ UNEP/SETAC project proposed)
- Test, refine
- Extend social collaboration
- Public broad announcement
- Long time maintenance and improvement



The "open source" part of the Registry

- Idea: Put no barriers for data providers to make data available
- Similar to wikipedia, low barriers can make existing knowledge more easily and faster available, and the result can be of high quality
- Also interesting playground (commenting on data sets, ...)
- Contains also several „liberated“, complete commercial data sets from ecoinvent
- Contribution of PRé Consultants to the Life Cycle Initiative



4. Spirit of the format converter and the registry



Spirit of the Registry, and of the Format Converter

- enabling exchange and collaboration;
- trying to solve practical problems: data availability, access, exchange, quality assessment, overcoming barriers;
- NOT try to outcompete others



5. Conclusions



Conclusions

- Both the Registry and the Format Converter 2.0 are close to a public release
 - for both, a community will be built, within two proposed Life Cycle initiative projects
 - Both registry and format converter help to work and "live" with diversity in LCA.
- This does not mean that they are positioned against efforts to harmonise. They only recognise currently existing diversity (and user needs).



Thank you!

Andreas Ciroth, Mark Goedkoop
ciroth@greendeltac.com, goedkoop@pre.nl

