CHALMERS

Quality assessment of the ecoinvent and SPINE@CPM databases based on the ISO/TS 14048 data documentation format

ANN-CHRISTIN PÅLSSON

IMI - Industrial Environmental Informatics
for
CPM - Centre for Environmental Assessment of Product and Material Systems
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2006

CPM Report 2006:12

Summary

The ISO/TS 14048 data documentation format describes and specifies relevant information that is needed when performing life cycle assessment in accordance with the ISO 14040 series of standards. This report describes how ISO/TS 14048 can be used as a practical tool to support data quality management in LCA and presents the results from a quality assessment of both the Swiss ecoinvent and the Swedish SPINE@CPM databases, based on requirements according to ISO/TS 14048 and the ISO 14040 series of standards.

The quality assessment includes three parts; a review of the quality routines that are applied, the data format that is used, and the content of the database in terms of documentation:

- *Quality routines:*
  Both databases have well-established routines for the management and administration, in terms of data acquisition, review, data publishing and technical administration. Both have specified quality requirements and guidelines to be applied for the database. The characteristics of the quality requirements as well as the general organization of data acquisition, however, differ between the two.
- *Data format:*
  The EcoSpold format is used for documentation of datasets in the ecoinvent database. This format is intended as a data exchange format, not as a data documentation format. The format is not easy to interpret and there are no manuals or guides available to facilitate interpretation, except for the technical specification. The format lacks structured fields for several parts of relevant information according to ISO/TS 14048, e.g. regarding system boundaries.
  The SPINE format, which is used in SPINE@CPM, is a data documentation format in a similar way as ISO/TS 14048. There are manuals and guides available to facilitate interpretation. The format lacks structured fields for some parts of information according to ISO/TS 14048, e.g. regarding validations that have been performed.
- *Database content:*
  The content of the ecoinvent database is distributed into two media; as EcoSpold formatted datasets and as separate reports describing datasets in more detail. The availability of information, based on ISO/TS 14048 and ISO 14044 requirements, varies depending on type of dataset. For “Unit process”, “Multi-output process” and “Impact category” datasets, the relevant information should be available when combining documentation in the EcoSpold format and in reports, but since it is distributed into two media it is not in an easily accessible form. The “Cumulative LCI result” and “LCIA result” datasets, however, lacks transparency and therefore interpretation in line with ISO 14044 is not possible for these datasets. For example, there is no information about which processes that is included in the calculation, system boundaries etc.
  The content of the SPINE@CPM database is documented in the SPINE format. There are no separate reports, but the documentation contains references to the sources that have been used in the compilation. The quality of the documentation of datasets varies, but sufficient or adequate information to assess different quality aspects according to ISO 14040 series should in general available.

Table of contents

Introduction

The ISO/TS 14048 data documentation format is designed to support and facilitate information management for LCA by transparent, reviewable and understandable documentation of data, which allows users to independently assess the quality of the data. The document describes and prescribes relevant information to be documented when performing life cycle assessment in accordance with the ISO 14040 series of standards. It supports and facilitates transparent documentation of data, by enabling that all relevant documentation for assessment and interpretation can accompany the data through all stages of data collection, storage and use. This facilitates transparency, review and accessibility of information, and supports assessment of data quality in line with quality requirements.

This report describes how ISO/TS 14048 can be used as a practical tool to support data quality management in LCA and presents the results from a quality assessment of both the Swiss ecoinvent database and the Swedish SPINE@CPM database, based on requirements according to ISO/TS 14048 and the ISO 14040 series of standards. The main focus of the quality assessment was the ecoinvent database, with comparisons of how the same issues are handled for the SPINE@CPM database.

The assessment has been performed within the CPM-project "Translating the ecoinvent database into SPINE and ISO/TS 14048 format". In this project, a software tool for automatic conversion of ecoinvent datasets into both the ISO/TS 14048 format and the SPINE format was developed. The purpose of the project was to make more data available for the partners of CPM and to promote the ISO/TS 14048 format. The results from this project is available at:
http://databases.imi.chalmers.se/imiportal/

The ISO/TS 14048 data documentation format

The ISO/TS 14048 data documentation format was published in 2002, and includes general and detailed requirements for data documentation of life cycle inventory (LCI) data, as well as requirements on formatting and reporting, i.e. requirements on how the data documentation format shall be used.

A model of a technical system, i.e. an LCI dataset, is in ISO/TS 14048 referred to as *Process*. The data documentation format for description of a process consists of three parts (see figure 1):

- *Process*, which contains the *process description* of properties of the modelled process with regard to technology, time-related and geographical coverage etc. The Process part also includes the *inputs and outputs* to the modelled process.
- *Modelling and validation*, which contains the description of prerequisites for the modelling and the validation of the process e.g. modelling choices describing which processes and flows that have been excluded.
- *Administrative information*, which contains general and administrative information related to the administration of the documentation of the process e.g. data commissioner, date completed, copyright etc.

For each part, detailed data fields are specified that holds specific information. The format can both be used to document individual unit processes and processes that consist of a combination of processes, e.g. an LCI product system.

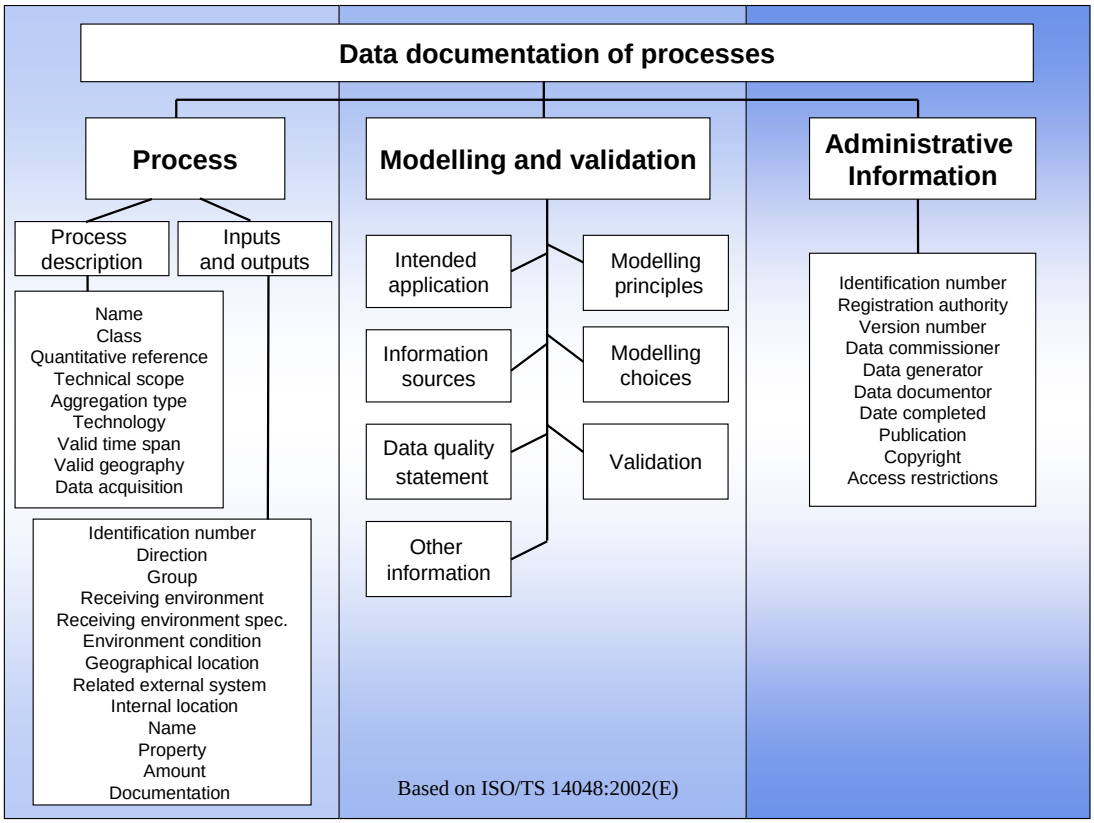

Figure 1. Illustration of structure of the major parts in the ISO/TS 14048 data documentation format (based on ISO/TS 14048), from Flemström & Pålsson (2003)

The ecoinvent database

The ecoinvent database contains more than 2700 datasets and is published via the Internet since 2003 (http://www.ecoinvent.ch). The datasets are delivered as:

- unit processes and multi-output processes (unit processes with more than one product)
- cumulative LCI results (which are calculated from the datasets included in ecoinvent)
- LCIA results (which are calculated from unit processes or cumulative LCI results together with impact assessment data that is included in the database).

That is, for each unit process cumulative LCI results as well as LCIA results can be calculated. The datasets can be viewed and downloaded from the database as EcoSpold formatted xml-files.

In addition, the database contains elementary flows datasets (specifies elementary flows that are documented in the LCI and LCIA datasets) and impact category datasets (impact assessment data that are used to calculate the LCIA results). These datasets may also be downloaded from the database as EcoSpold formatted xml-files. The database is also sold together with several commercial LCA software packages.

Ecoinvent contains data for a large variety of different processes and industrial sectors, and are compiled by the different research institutes participating in the ecoinvent center. The database contains data for

- Energy: Energy supply, Fuels, Heat production and Electricity production

- Materials: Plastics, Paper and Board, Basic Chemicals, Detergents, Metals, Wood
- Building materials
- Agricultural products and processes
- Transport services
- Waste treatment services

The SPINE@CPM database

The SPINE@CPM database contains more that 500 datasets and is published via the Internet since 1998 (http://databases.imi.chalmers.se/imiportal, or http://www.globalspine.com). The datasets is documented as unit processes or as composite systems, i.e. as a process that consists of a combination of documented unit processes. Datasets are delivered as SPINE or ISO/TS 14048 formatted reports (rtf-, pdf-, or html-files) or communication files (xfr- or xml-files).

The database also contains impact assessment data for three impact assessment methods documented in the IA98 data documentation format, which is a format designed for transparent documentation of impact assessment data, in line with requirements in ISO 14044. The impact assessment datasets are freely available and can be viewed in a web-tool named WWLCAW and downloaded in full in a MS Access database file.

SPINE@CPM contains a variety of different processes, and have been acquired and documented either by the partners of CPM, i.e. by the industrial partners and by Chalmers University of Technology, or through data exchange agreements with e.g. M.Sc. students performing LCA:s. The database contains datasets describing:

- Different modes of freight transports (road, rail, air and sea)
- Energyware production (electricity, heat and fuels)
- Production of selected materials (e.g. chemicals, natural materials, polymers, fertilizers, metals, building materials, road material, electronic component groups and steel products)
- Waste management

Using ISO/TS 14048 for data quality assessment

A central element when performing and interpreting life cycle assessment (LCA) studies is the specification of data quality requirements and the assessment of the data quality in line with the requirements. Data quality is in ISO 14040 generally defined as "*characteristics of data that relate to their ability to satisfy stated requirements*". The ISO 14044 standard specifies a number of aspects for which data quality requirements should be set when performing LCA:

- time-related, geographical and technology coverage,
- precision,
- completeness,
- representativeness,
- consistency,
- reproducibility,
- sources of the data, and
- uncertainty of the information.

Transparency is furthermore stated as a fundamental principle when conducting LCA:s in ISO 14044.

Consequently, transparent information to assess the different quality characteristics and requirements should be available when performing LCA. The quality of any given life cycle inventory (LCI) dataset in a specific study may only be determined through a thorough knowledge of the described system and of the data. Therefore, documentation of the how the dataset was acquired and compiled is important to ensure that the data can be correctly used and interpreted.

The ISO/TS 14048 format supports data quality management in LCA by specifying relevant information that should be available to enable assessment and interpretation of both qualitative and quantitative aspects of data quality for life cycle inventory datasets and studies, in accordance with ISO 14044. The fundamental LCA principle of transparency also implies that availability of documentation of data can be considered as a quality aspect in itself.

With regard to quality aspects that should be considered according to ISO 14044, the ISO/TS 14048 format specifies relevant information with regard to technology, valid time span and geography, as well as relevant information to both express and describe precision and uncertainty of the data. The format also specifies relevant information to describe and specify the sources that have been used, as well as the principles used and the choices made in the modelling. This information is important for the reproducibility of the information. The overall information specified by the format can be used to assess and determine completeness, representativeness and consistency. The format also specifies how processes that consist of a combination of documented unit processes (e.g. LCI product systems) can be transparently described. Figure 2 provides an overview of how the different quality aspects that should be addressed in LCA studies can be interpreted in terms of the different parts of the ISO/TS 14048 format. The figure is, however, not intended to be exhaustive or complete with regard to how these data quality aspects are addressed in ISO/TS 14048.

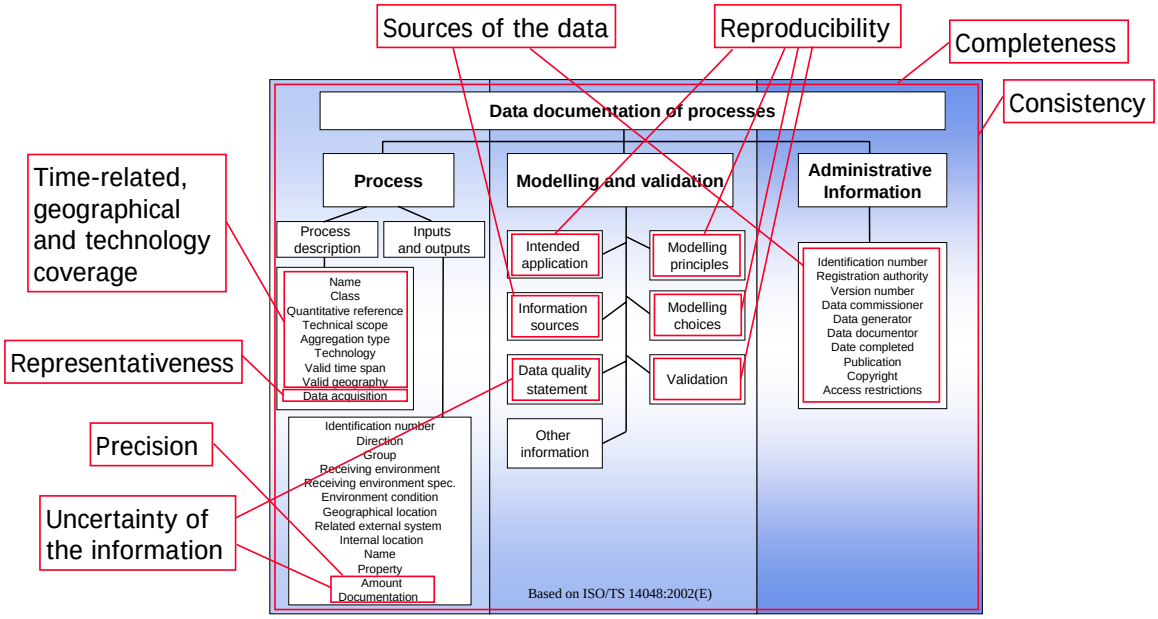

Figure 2. Overview of how different quality requirements, which should be addressed according to ISO 14044, can be interpreted in ISO/TS 14048.

Thus, the ISO/TS 14048 format specifies relevant information to consider and assess when performing LCA. In this way, the format can be used as a checklist during the goal and scope phase, to specify detailed quality requirements for both the study and for individual included unit processes. Well specified and detailed quality requirement facilitates data quality assessment. The format can be used during data collection and interpretation to determine and assess if relevant information to assess quality requirements is available, as well as a basis to assess if and how specified quality requirements are fulfilled. This provides both a consistent knowledge of the included unit processes and of the product system, and a structured work procedure for information management during LCA projects. Missing information and knowledge gaps can easily be identified and highlighted. ISO/TS 14048 can be applied as a checklist on any data source or any database, irrespective of original format for data storage and communication.

Purpose and scope of quality assessment of ecoinvent and SPINE@CPM databases

A quality assessment has been performed of the ecoinvent and the SPINE@CPM LCA databases, based on the ISO/TS 14048 data documentation format. The purpose of the quality assessment is to assess and benchmark whether sufficient information is available to allow for interpretation and assessment of different data quality aspects according to ISO 14044. For this assessment, the ISO/TS 14048 is used as a basis to determine if the relevant information to assess data quality aspects in line with requirements and recommendations in ISO 14044 is available in the two databases. The main focus of the assessment is the ecoinvent database, with comparisons with how these issues are addressed in the SPINE@CPM database.

The quality assessment includes three parts:

- *Quality routines* that are applied for the databases, with regard to quality requirements, review, maintenance and support
- *Data format* in which datasets is published and delivered
- *Database content* in terms of documentation of datasets, with regard to transparency, interpretability and availability of information.

The scope for ISO/TS 14048 is life cycle inventory. Both databases, however, also contain impact assessment data, and therefore a brief assessment of the impact assessment datasets in the databases is also included, with regard to data format, compatibility and transparency.

The assessment does not include any evaluation of the representativeness of datasets for real processes and market situation or the quality of the information sources used in the modelling of datasets.

The assessment of ecoinvent is based on the reports available from the ecoinvent network and on random chosen individual datasets from the database. The assessment of SPINE@CPM is based both on reports and experiences from work with the database.

Note: The assessment was performed within the research group Industrial Environmental Informatics (IMI). This research group is also responsible maintenance of the SPINE@CPM database. The author of this report has been actively involved with the establishment and maintenance of this database, and has both been responsible for developing and implementing quality routines as well as for the everyday management of it. The author was also actively involved in both the national and international working group for the development of ISO/TS 14048. Thus, the assessment may be considered to be biased. The author has, however, made efforts to report results as objectively as possible.

Results from quality assessment

The following sections contain the results from the quality assessment of the Swiss ecoinvent and the Swedish SPINE@CPM databases in terms of:

- Quality routines
- Data format
- Database content

The results from each part are reported through a brief introduction, and then the results are described for the two databases respectively.

Quality routines

Both the ecoinvent and SPINE@CPM have specified quality requirements and guidelines that are applied, and have well-established routines for the administration and maintenance of the database, in terms of data acquisition, data review, data publication and technical administration (see figure 3 below).

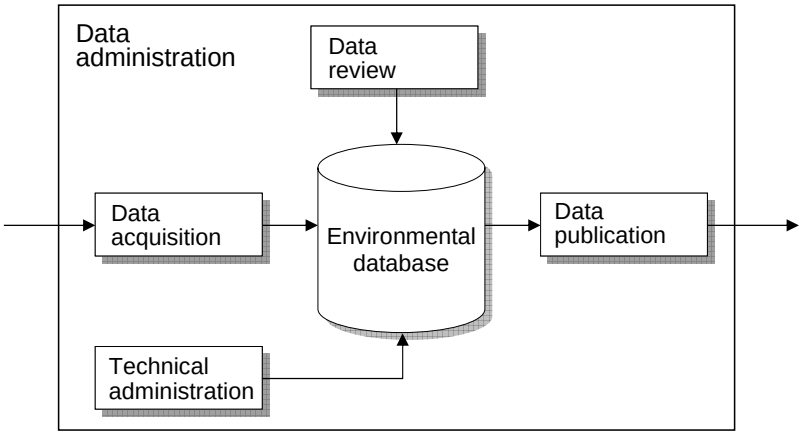

Figure 3. General organisational functions for an environmental information system, based on Carlson & Pålsson (1998)

The characteristics of the quality requirements that are applied, however, differs between the two databases. For the ecoinvent database, there are detailed guidelines that should be applied in the modelling and compilation of datasets, which are aimed to harmonise the content of the database. There are also requirements on documentation. For the SPINE@CPM database, the main requirement is structured documentation in terms of detailed specifications on information that should be available for datasets, to allow for interpretation of choices made in the modelling and compilation of datasets.

The data acquisition procedures also differ between the two databases. For ecoinvent, the participating institutes in the ecoinvent center contribute with data through data acquisition projects aimed to acquire datasets for e.g. specific materials. Thus, there is an ambition to be comprehensive in terms of coverage of different industrial sectors. For SPINE@CPM, the data acquisition is performed through “in kind” efforts of the partners of CPM or through data exchange agreements. The datasets are acquired in different internal projects for different purposes, which results in a wide variety of different processes in the database, but also that the coverage in terms of industrial sectors are scattered.

Below follows an overview of the quality routines that are applied for the ecoinvent and SPINE@CPM databases, respectively, in terms of maintenance, quality requirements or guidelines, data acquisition, review, corrections, and support.

ecoinvent

The ecoinvent quality routines include both specific guidelines that should be followed when compiling datasets for ecoinvent, and an internal review procedure that is performed before data is entered into the ecoinvent central database (see figure 4). Also, there is a procedure to release corrections and updates of data. Some limited support on the actual content of the datasets is also available.

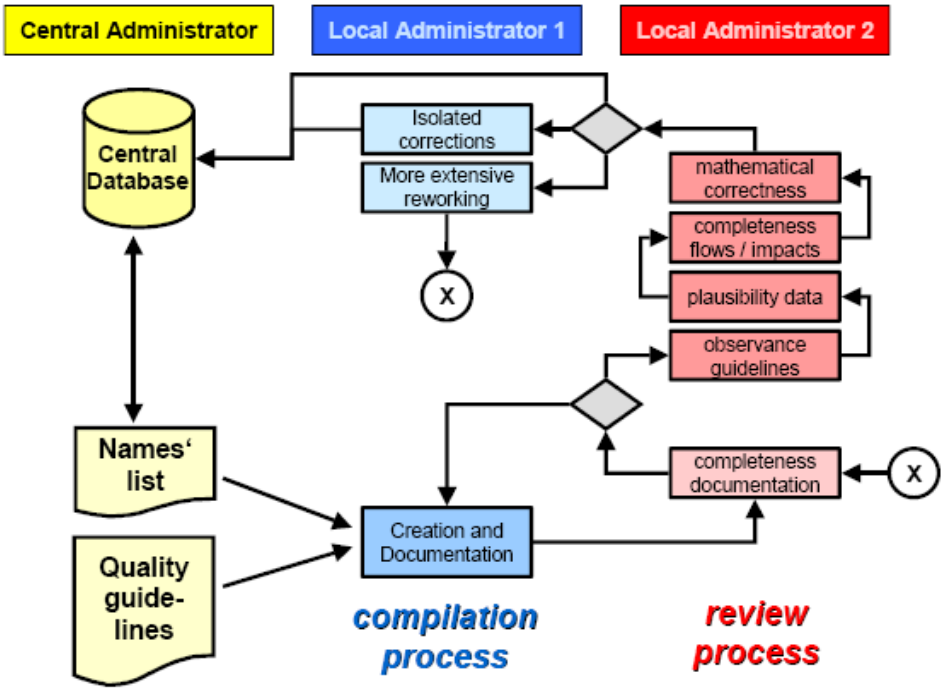

Figure 4. Overview of the internal review and data quality control within the ecoinvent project , from Frischknecht et al (2004)

Maintenance

The ecoinvent database is maintained by the ecoinvent center, also known as the Swiss Centre for Life Cycle Inventories, which is a joint initiative of institutes and departments of the Swiss Federal Institutes of Technology Zürich (ETH Zurich) and Lausanne (EPFL), of the Paul Scherrer Institute (PSI), of the Swiss Federal Laboratories for Materials Testing and Research (EMPA), and of the Swiss Federal Research Station Agroscope Reckenholz-Tänikon (ART). The ecoinvent Centre was and is supported by Swiss Federal Offices.

The responsibility for the management of the ecoinvent center is held by EMPA, one of the participating institutes.

Quality guidelines for data collection and inventory

The ecoinvent network has developed guidelines for how to model and inventory the datasets to be published in the database. The guidelines are described in Frischknecht

et al. (2004), and they provide specific details on how to model and inventory datasets with regard to:

- *Temporal, geographical and technical scope*
- *Modelling principles* with regard to aggregation level of processes, cut off rules, assumptions on how to handle missing information, market situation, electricity mixes, etc..
- *Naming rules for products and services*
- *Documentation of elementary flows* with regards to level of detail, naming rules, units, as well as rules for specific airborne, water and soil pollutants and land occupation and transformation. Also, some specification of neglected elementary flows is provided.
- *Allocation rules* for multi-output processes
- *Estimation of uncertainties*.

Parts of the guidelines are general, and they do not give any specific recommendations regarding e.g. cut-offs or allocations to be applied. Consequently, the methodological choices for different datasets will be different depending on the knowledge and experience of the person modelling and collecting the data.

Below is an overview of the content of the different parts of the quality guidelines, based on Frischknecht et al (2004).

Temporal, geographical and technical scope

The datasets in ecoinvent is collected for the market and consumption situation in Switzerland for the year 2000. In most cases the processes represent the average of currently used technology. Emissions from past, present and future are included.

Modelling principles

Aggregation: The ambition for unit processes is to present data at the "smallest" level at which data is available, and that these unit processes can be used to model different LCA product systems. In Frischknecht et al. (2004) it is stated that the database contains data that are neither vertically nor horizontally aggregated, but it is also stated that average data for a country or a region are calculated with available data from different suppliers if they use comparable processes. This means that these processes in fact are horizontally aggregated.

Cut-off rules: No strict cut-off rule is followed in the ecoinvent database. It has been up to the environmental knowledge of the person inventorying the data to decide whether to include the production of a certain input or the release of a certain pollutant. Also, although not stated explicitly, if the production of an input material is excluded or not available in the ecoinvent database, then the amount of the input material is also excluded from the inventory. Thus, there may be data gaps in terms of inputs of materials for individual datasets.

Handling of missing information: Regarding missing information; if the data availability is poor, stoichiometric balances is used to determine the raw material demand, with the assumption of 95 % yield. For the releases an educated guess is made based on plausibility considerations. If such assumptions dominated the LCI results, further investigation is made. Otherwise it is kept in the inventory.

Processes and products modelled as services: Both transports and waste treatment are modelled as services delivered to the production processes. Transportation work are recorded as inputs of the total amount of transport-work for all different materials etc. used in the production process, by different means of transport. It is not possible to distinguish the specific transport-work for a specific material. Waste treatment is recorded as an input of the total amount of waste in the production process that is treated by different waste management processes. This way of handling transports and waste management may be confusing for LCA practitioners who are used to model waste management and transports as separate processes in the same way as other processes.

Infrastructure, e.g. manufacturing of cars, erection of power plants, construction of roads, etc, is also recorded as an input in inventories for production processes. Land occupation and transformation of a production process is normally inventoried for infrastructure processes.

Naming rules

Products and services are named according to the following rules:

- Name of product/service, production process, and treated product, and treatment level. Names of chemicals can contain the sum formula and their weight-% in water or the respective carrier material.
- Additional information, e.g. description of product preparation, system boundary, etc
- Destination ("at", "to" or "in") and level of value chain (e.g. factory gate)

Thus, the name of the dataset contains a lot of information, some of which is also documented in different field in the EcoSpold format.

The name of the product or service is also used for naming the process delivering the product or service, i.e. the LCI dataset. For example, the dataset named "sheet rolling, aluminium" delivers the reference product "sheet rolling, aluminium". Consequently, inputs of materials to production processes are recorded using the same name as the process delivering the material. Thus, no clear distinction is made between product name and corresponding process name.

Dissolved chemicals are quantified in terms of the pure active substance, e.g. 1 kg "sodium hydroxide, 50% in H2O" refers to 2 kg NaOH with water content of 50%.

Documentation of elementary flows

Elementary flows are registered at the most detailed level for which information is available. The naming rules are based on the work in SETAC working group on "Data availability and data quality". The environmental compartment is described by a category (air, water, soil, and resource) with a further specification by a subcategory (e.g. "low population density", "lake", etc.).

Additionally, specific guidelines are given as to how to report certain airborne, water and soil pollutants as well as resource use. The following elementary flows have been neglected: noise emissions, water vapour, nitrogen and oxygen releases to air, as well as water emissions to air, soil and water (except for emissions of water from air planes).

Allocation rules

The ecoinvent database contains both unallocated multi-output processes, as well as their derived allocated processes resulting from applying allocation factors on an unallocated process. The allocation factors that are applied are documented in the multi-output process, and may also include a description of the allocation method. The sum of all allocation factors for one specific flow is always 100%. System expansion is avoided whenever possible.

There is, however, no specification of a recommended allocation method to be applied. The methods specified by the EcoSpold format are “undefined”, “physical causality”, “economic causality” and “other method”. Also, different allocation methods could in principle be used for different flows within the same dataset.

Estimation of uncertainties

All LCI datasets in ecoinvent is supplied with uncertainty information according to lognormal distribution. However, the uncertainty information (standard deviation) is estimated, since the amount for a specific input or output in a dataset has generally only been available as a mean value.

The estimation of uncertainty information is made with a standard procedure, where the standard deviation for inputs and outputs is calculated by means of a qualitative assessment of quality according to a pedigree matrix, combined with basic uncertainty factors. For cumulative LCI results, the standard deviation is calculated by Monte-Carlo simulation.

The assessment of data quality according to the pedigree matrix is made with regard to six different indicators, where a score is assigned to each of the indicators. The indicators are “Reliability”, “Completeness”, “Temporal correlation”, “Geographical correlation”, “Further technological correlation” and “Sample size”. Default uncertainty factors are used for each score, and the scores that are assigned for the datasets are used in the calculation of standard deviation for the flows.

The basic uncertainty factors depend on the kind of input or output that are considered. They are based on general assumptions that e.g. CO2 emissions generally show a lower uncertainty than CO emissions. These factors are based on expert judgement. The basic uncertainty factor is also included in the calculation of standard deviation.

It is stated that for some datasets, slightly different approaches have been used for uncertainty assessment, which is described in the respective ecoinvent reports (Frischknecht et al. 2004). Also, it is stated that statistical methods has been used when enough samples were available.

The fact that the uncertainties generally are estimated, and that the estimation is based on several subjective judgments, can make the uncertainty information difficult to use and interpret.

Data acquisition

The data in the ecoinvent database have been acquired by the research institutes participating in the ecoinvent center. The tasks are distributed according to the expert knowledge of the partners. See table 1 for an overview of the institutes responsible for specific parts of the database content. The data acquisition is performed within projects aimed to acquire datasets for e.g. specific material groups. There is a plan to release a new version of ecoinvent during the year 2007, which will contain data for new industrial sectors and more data for the existing sectors in the database.

Table 1. Database content, responsible institute and their partners in LCI data compilation (from Frischknecht et al 2004)

Database content	Responsible Institute	Partners
Energy supply Fuels Heat production Electricity production	Paul Scherrer Institute	ESU-services
Plastics Paper and Board Basic Chemicals (I) Detergents Waste treatment services	Swiss Federal Laboratories for Materials Testing and Research (EMPA), in St. Gallen	Doka Life Cycle Assess-ments, Chudacoff Ökoscience
Metals Wood Building materials Basic chemicals (II)	Swiss Federal Laboratories for Materials Testing and Research (EMPA), in Dübendorf	Chudacoff Ökoscience
Transport services	Natural and Social Science Interface, Swiss Federal Institute of Technology Zurich (ETHZ)	
Basic chemicals (III)	Institute for Chemical and Bioengineering, Safety and Environmental Technology Group, Swiss Federal Institute of Technology Zurich (ETHZ)	
Agricultural products and processes	Swiss Federal Research Station for Agroecology and Agriculture (FAL)	Swiss Federal Research Station for Agricultural Economics and Engineering (FAT)

Review

Before a dataset is entered into the central ecoinvent database, an internal review process is performed. The review is performed by someone at a different institute within the ecoinvent network and concerns the following issues:

- Completeness of the documentation. Checks that all investigated datasets are described in the report, and that meta information and flow data is available for each dataset.
- Consistency with the quality guidelines. Checks whether the unit processes have been modelled according to the ecoinvent quality guidelines.
- Plausibility check for the life cycle inventory data. Checks selected input and output flows for plausibility.
- Completeness of inputs and outputs. Control of completeness of flows based on the environmental and technical knowledge of the reviewing person.

Reviewers are not necessarily technical experts of the processes reviewed. They were supported by the person responsible for the report in such cases.

- Mathematical correctness of calculations. Control of selected inputs and outputs for mathematical correctness, e.g. the transport service inputs, the waste heat or CO2 emissions.

After review, the reviewer gives feedback to the person responsible for the dataset. The feedback includes a proposal for the further procedure:
• dataset import in database without changes
• dataset import in database after (minor) corrections of the data
• dataset import in database after a major revision of the data and after a second review

The role of the review is to help improve the quality, but it is up to the person responsible for the dataset whether or not to consider all proposals for corrections of data. Thus, the sole responsibility for the contents of datasets remains with the person and institute who investigated the data.

According to Frischknecht et al. (2004), a validation comment is written in the documentation of the dataset by the person responsible for the review before import of a specific dataset into the central database. However, a review of a number of random chosen datasets showed that the comment from the internal review only contains the statement “Passed”, and not any further information results from the review, e.g. issues that were identified during the review and whether they have been corrected.

Corrections

The ecoinvent network has a procedure for correcting errors found in the published datasets. Each dataset has a version number, and discovered errors are corrected in a next version. Since the release of the database in 2003, three new versions have been published. All earlier versions of the database are available. The members of the ecoinvent database are informed about errors that have been discovered and its expected date of correction via the Internet (through the ecoinvent website), and via e-mail (in particularly severe cases only).

Support

According to the “Terms of Use”, standard support includes intermediate updates of the ecoinvent Data (error corrections), access authorisation to the ecoinvent mailing list, to the solution collection "frequently asked questions (FAQ)", and the possibility of direct contact with the licensor (the ecoinvent centre or the reseller) via e-mail or fax. Questions are answered by e-mail or fax, and the stated response time is typically a maximum of five working days.

SPINE@CPM

The quality routines for the SPINE@CPM database include both quality requirements for data documentation and a review which is performed to check if the requirements are fulfilled. There is also a procedure to release corrections of published data. Limited support on the content of the database is available.

Maintenance

The SPINE@CPM database is maintained and managed by the research group Industrial Environmental Informatics at Chalmers University of Technology, on commission from CPM (Center for Environmental Assessment of Product and Material Systems).

CPM is a Swedish national competence center, financed by VINNOVA (Swedish Governmental Agency for Innovation Systems), Chalmers University of Technology and Swedish industry. Research and development within the center is performed in close collaboration between academy and industry. Since the start of CPM in 1996 the following companies have participated: ABB, Akzo Nobel, Avesta Sheffield, Bombardier, Cementa, Duni, Electrolux, Ericsson, IKEA, ITT Flygt, M-real (MoDo), Norsk Hydro, Perstorp, SAAB Automobiles, SCA, Stora Enso, SwedPower (Vattenfall), Telia Sonera (Telia), Tetra Pak, AB Volvo, Volvo Cars.

Quality requirements

The quality requirements for the SPINE@CPM database are formulated as documentation requirements, i.e. as required information that needs to be available for each dataset in the database (Arvidsson et al 1997 and Arvidsson et al 1999). The requirements are expressed in terms of the quality aspects specified in the ISO 14040 series of standards. The general idea is that the documentation should to enable for the different users of the database to assess and decide the quality and applicability of a certain dataset for the specific application in which he or she intends to use it. In this sense data quality is considered to correspond to documentation quality. If the data is not documented, or the documentation is difficult to interpret, it is impossible to assess any other aspect of data quality.

The quality requirements was published in 1997 and were developed during the establishment of SPINE@CPM, in a consensus-process where the following industrial CPM partners participated: ABB, Akzo Nobel Surface Chemistry, Perstorp, SCA, Stora (now Stora Enso), Vattenfall and AB Volvo.

In short, the documentation of models of technical systems (i.e. LCI datasets) according to the CPM quality requirements consists of six separate, but closely integrated sections, see figure 5. The type of information within each section regards different choices made during modelling and data acquisition. To fulfil the requirements, all sections should be documented. The model is described in qualitative terms regarding how the model was the designed, together with the quantitative data for input and output flows and information regarding how the numerical data was obtained. Also, more general aspects are described, for example how the model should be used and how the information may be distributed.

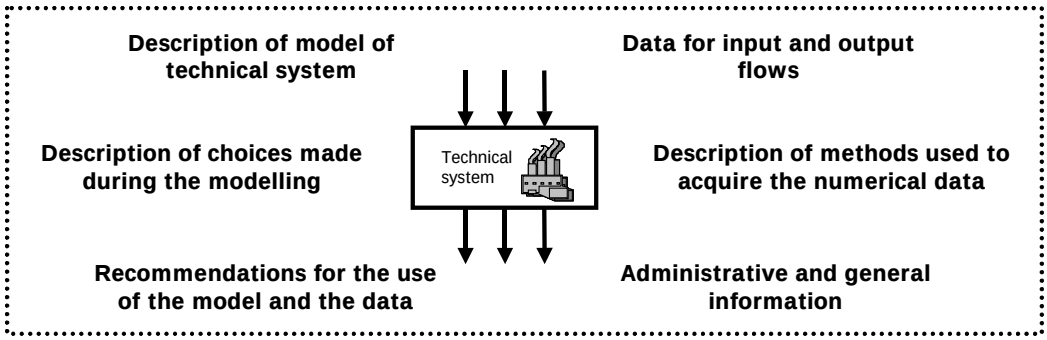

Figure 5. The CPM data documentation criteria (in short), based on Pålsson (1999a)

For each section, detailed information items (data fields) are specified in terms of the SPINE and the ISO/TS 14048 data documentation formats. The different types of information are:

1. *Description of the model of the technical system*, where the content of the system is described, with regards to, for example, the included processes etc.
2. *Description of choices made during the modelling* and the objective for the choices, e.g. the purpose of the data acquisition, the choice of functional unit and system boundaries etc.
3. *Data for input and output flows* to the system, i.e. quantitative data on flows together with information on the origin and destination of the flows in terms of environmental media and type of flow.
4. *Description of methods used to acquire the numerical data*, i.e. a transparent description of the assumptions and sources used for the quantitative data.
5. *Description of recommendations for the use of the model and the data*, e.g. areas of application and a general description of the quality of the sources that have been used.
6. *Administrative and general information* regarding, for example, how data may be distributed and who is responsible for acquiring and documenting the data.

With regard to further guidelines for the collection and compilation of datasets for SPINE@CPM, there are some guidance for documentation of elementary flows (Erixon et al 2000), but there are no specific guidance with regard to how the datasets should be modeled in regards to e.g. cut-offs and allocations. Consequently, the datasets in the database are differently modeled depending on the original purpose for which the data was acquired. There are recommendations to make data available at the most detailed level at which data is available, e.g. that LCI results are documented through transparent LCI flow charts where each of the included processes are documented separately, but it is not a requirement.

Data acquisition

The data published in SPINE@CPM are acquired by the partners of CPM or by data exchange agreements with e.g. other organisations or M.Sc thesis students performing LCA:s. The data in the database has generally been acquired within different LCA projects, and have then been documented according to the quality requirements and submitted for publishing in SPINE@CPM. There has also been some specific data acquisition projects aimed to acquire prioritised general data, such as transport and energyware production.

Review

The general review procedure at SPINE@CPM is shown in figure 6. The review for SPINE@CPM is performed by the data administration at Industrial Environmental Informatics at Chalmers University of Technology and concerns (Pålsson 1999b):

- completeness, coherency and consistency of the documentation, i.e. that the information is sufficient to be able to assess and determine the quality and applicability of the dataset
- that the information is understandable
- that the documentation format is used correctly
- that the information is relevant for the use of the data

There is, however, generally no review of e.g. the agreement of the information with the original source or plausibility checks. This type of review has only been performed for some specific datasets, e.g. for some electricity and transport datasets, where the original generator of data has checked the datasets.

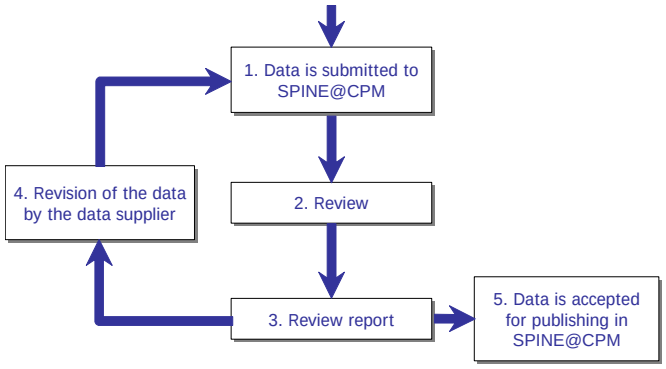

Figure 6. Data review procedure at SPINE@CPM, from Pålsson (1999b)

Datasets are not published until they have passed the review, but there is no specific information from the result from the review available, except for a specification of who has performed the review and a classification of the quality of documentation. The classification is made in three categories; Sufficient (i.e. sufficient information to use the data should be available), Acceptable (i.e. the information may be sufficient, but will most likely require further investigation by the data user to be able to use it) and Unsatisfying (i.e. the information is not sufficient and will require further information by the user to be able to use it). Generally, datasets that are classified to be unsatisfying is not published in the database. They are instead revised and supplemented by the data supplier, until they can be classified as sufficient or acceptable.

Corrections

Users of the database are encouraged to report errors that have been discovered in datasets to the data administration at Chalmers. The reported errors are forwarded to the person responsible for the dataset for correction. After correction, the corrected dataset is published together with documentation in the dataset of changes that have been made since the dataset was first published. Earlier erroneous versions of datasets are not available.

Support

Limited support is available for the database, where users can contact the data administration at Chalmers via e-mail. In general, the support concerns how to interpret datasets and how to get access to data.

Data format

The data format used for storage, publication and communication of data in the database has a large influence on the availability of information, i.e. whether the data format allows the relevant information to be stored, presented and communicated. In the ecoinvent database the EcoSpold format is used for this purpose. The SPINE@CPM database is available in both the SPINE format and the ISO/TS 14048 format, but the datasets are originally stored in the SPINE format and therefore only the SPINE format is discussed. Datasets available in the ISO/TS 14048 format has been automatically translated from SPINE into ISO/TS 14048.

Both the EcoSpold and the SPINE format lack structured fields for parts of the information that is needed according to the ISO/TS 14048 format. The EcoSpold format, however, lacks more information compared to the SPINE format, possibly due to the fact that the EcoSpold format is intended as a data exchange format, whereas SPINE is intended as a data documentation format in the same way as ISO/TS 14048.

Note: ecoinvent datasets can also be translated into the ISO/TS 14048 format. A software tool is available that automatically converts ecoinvent EcoSpold datasets into ISO/TS 14048. The converted datasets can be downloaded as printable Adobe Acrobate pdf-reports, as HTML reports and as xml-files formatted in the ISO/TS 14048 format. The detailed specification of the translation between the formats is available in Erlandsson et al (2006). The tool was developed by IMI, on commission of CPM, and is freely available at: http://databases.imi.chalmers.se/imiportal

ecoinvent

All datasets in the ecoinvent database is delivered in the EcoSpold data format. The format consists of two parts; Meta information and Flow data (see figure 7). The general content and structure of the two parts is outlined in Table 2.

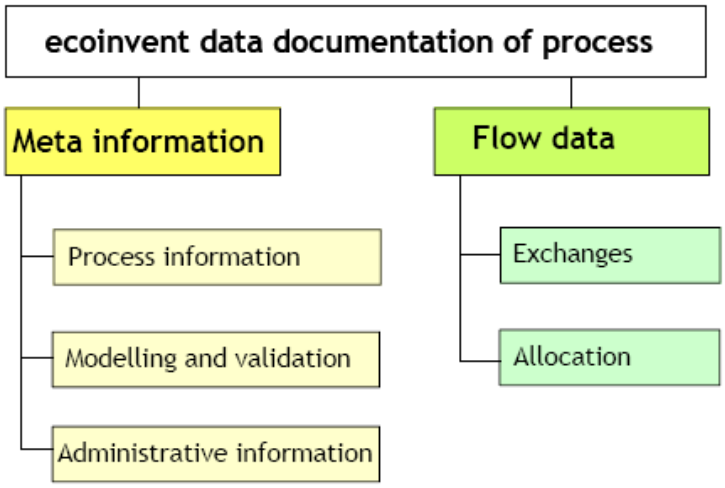

Figure 7. Overview of the structure of the EcoSpold format, from Frischknecht et al. (2004)

Table 2. The general structure of the EcoSpold format, from Frischknecht et al. (2004)

Meta information	
Process	
ReferenceFunction	defines the product or service output to which all emissions and requirements are referred
TimePeriod	defines the temporal validity of the dataset
Geography	defines the geographical validity of the dataset
Technology	describes the technology(ies) of the process
DataSetInformation	defines the kind of process or product system, and the version number of the dataset
Modelling and validation	
Representativeness	defines the representativeness of the data used
Sources	lists the literature and publications used
Validations	lists the reviewers and their comments
Administrative information	
DataEntryBy	documents the person in charge of implementing the dataset in the database
DataGenerator AndPublication	documents the originator and the published source of the dataset
Persons	lists complete addresses of all persons mentioned in a dataset
Flow data	
Exchanges	quantifies all flows from technical systems and nature to the process and from the process to nature and to other technical systems
Allocations	describes allocation procedures and quantifies allocation factors, required for multi-function processes

The EcoSpold format allows for documentation about the process and the inputs and outputs, but the format lacks structured fields for several parts of information according to ISO/TS 14048. For example, there are no specific fields to document system boundaries (towards other technical systems and towards the environmental system), the purpose and the applicability of the dataset, data treatment, etc. Also, the format does not allow for documentation of transparent LCI product systems, where each of the included processes is documented separately.

EcoSpold is referred to as a data exchange format in the technical documentation of ecoinvent (Hedeman & König 2003), and it seems that it is not intended as a documentation format, in the same way as ISO/TS 14048 or SPINE. Some parts of the format are very detailed in regards to specific information, whereas other parts of the format are very general.

The format includes several fields for indicator types of information. For example separate fields are used to distinguish between processes that are classified as infrastructure processes and other processes, and to indicate whether infrastructure are inventoried for the unit process or included in the calculation of cumulative LCI results. Coded nomenclatures with predefined options are used in several parts of the format, i.e. the choice is indicated by a code. For example, data field “energyValues” in Dataset information has the options 0=Undefined and 1=Gross values.

Some fields are used to for very different types of information. For example the generalComment field in the Reference function part is defined as to include information about the general application of the dataset, the information sources used in the compilation, the principles used when selecting which data to include as well as

choices made in the modelling. In the actual datasets in the ecoinvent database, this field may also include information about allocations that have been performed and further information about the process that is described.

It should be noted that the EcoSpold format makes use of ISO/TS 14048 terminology in terms of the concepts: Process, Modelling and validation and Administrative information (see figure 7). However the meaning of these concepts in the EcoSpold format is slightly different. For example, the process information part of EcoSpold is a mixture of product and process information, as well as modelling and validation information, whereas the process information part of ISO/TS 14048 is only intended as a description of the process.

With regard to support for interpretation and use of the format, there are no specific manuals or guides available, except for the technical documentation of the format which provides a specification of the format. The technical specification provides brief descriptions of the content of the fields but is somewhat difficult to use, ambiguous in some parts, and requires some effort to get acquainted to.

Also, the scope of the EcoSpold format is not well defined. The same data fields in the format is used for all types of datasets in the ecoinvent database; i.e. for life cycle inventory (LCI) data, life cycle impact assessment (LCIA) results, impact assessment data and elementary flows. Consequently, the different fields in the format are interpreted differently depending on which type of dataset that is documented. The technical specification of how the fields in the format should be interpreted for the different types of datasets is not easily understood for datasets other than LCI data. This leads to difficulties in interpreting the information. The original purpose of EcoSpold is LCI datasets, and the format is best described for this purpose. The specification for impact assessment data and elementary flows is, however, not well described and therefore the interpretation of the format for these types of datasets needs to be derived from a review of how the format is actually used in the database.

There are two very central fields for the interpretation of which type of dataset that is documented in the format: the fields Type (201) and impactAssessmentResult (208) in Dataset Information. These two fields together supply the key to what type of dataset that is documented in the format, and consequently to how all other fields in the format should be interpreted (as LCI data, as IA results, as elementary flows, or as impact categories). They are used to distinguish different types of datasets from each other. This important fact is however not clear from either the supporting material for the database or in the specification of the format.

With regard to impact assessment data, the EcoSpold format lacks a specification of the main parts of impact assessment based on ISO 14044. For example, the format lacks a specification for documentation of category indicators and provides no clear distinction between characterisation and weighting.

SPINE@CPM

The SPINE@CPM database is published and delivered in both the SPINE and the ISO/TS 14048 formats. All datasets have, however, originally been documented in the SPINE format, and have then been automatically translated into the ISO/TS 14048 format.

The SPINE format is a data documentation format, in a similar way as the ISO/TS 14048. The central concepts for LCI-datasets and flow charts in SPINE are 'Activity' and 'Flow' (see figure 7).

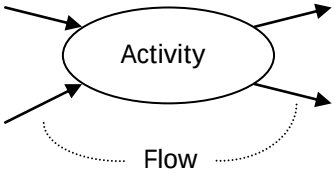

Figure 8. The concepts activity and flow in SPINE, from Carlson et al (1995)

An 'Activity' corresponds to the concept Process in ISO/TS 14048. An 'Activity' consists of the concepts 'ObjectOfStudy', 'Inventory', 'Flow' and 'QMetaData'. The definition of each of these concepts is described in table 3.

Table 3. The concepts in SPINE describing an LCI dataset

Activity	
ObjectOfStudy	Description of the technical system
Inventory	Description of choices made in the data acquisition and the objective for the choices (e.g. choice of functional unit and system boundaries etc.), Recommendations for the use of dataset General and administrative information
Flow	Input and output flows to the technical system
QMetaData	Description of the methods used to acquire the numerical data for inputs and outputs

The format allows for documentation about the process and the data, but the format lacks structured fields for some parts of information according to ISO/TS 14048, such as description of any validations that have been performed of the data, as well as possibility to express uncertainty information in terms of different distribution functions. The SPINE format includes different qualitative text based information, which allows most of the information according to ISO/TS 14048 to be documented, but is less detailed with regard to specific indicator types of information.

The format also allows for transparent documentation of processes that consists of a combination of unit processes, in a similar way as in ISO/TS 14048. Thus, it is possible to document LCI flow charts transparently, where each of the included processes is documented separately.

With regard to support and guidance for the use of the format, there are manuals available that facilitate documentation and interpretation of data documented in the

format, see for example Pålsson (1999a) and Flemström&Pålsson (2003), in addition to the technical specification of the format (Carlson et al 1995).

The impact assessment data in SPINE@CPM is documented using the IA98 data documentation format. The IA 98 is a data documentation format that is developed to support transparent documentation of impact assessment methods and data, based on requirements and recommendations in ISO 14044 (Carlson&Steen 1998, Carlson&Pålsson 2002). This format distinguishes and structures documentation for the different parts of impact assessment; impact categories, category indicators, characterisation and weighting, as well as the combination of the different parts into a full impact assessment method.

Database content

The datasets in the databases are documented and published in their respective data formats. In ecoinvent, the datasets are also further documented in separate reports that are referenced in the EcoSpold documented datasets. In SPINE@CPM all documentation of a dataset is available and documented using the SPINE-format, and is also available in the ISO/TS 14048 format. The general ambition for both databases is to provide users with transparent datasets to be used in LCA studies. However, the transparency of the different types of datasets that are published in the ecoinvent database varies.

Below is an overview of the content of the two databases with regard to transparency, interpretability and accessibility of information, based on requirements according to ISO/TS 14048 and ISO 14044.

ecoinvent

The ecoinvent database contains more than 2700 datasets documented as unit processes. The data in ecoinvent are implemented by documentation in datasets in the EcoSpold format, together with reports describing the datasets in more detail. The content, in terms of available documentation of data varies depending on type of dataset, i.e. Unit process, Cumulative LCI result, LCIA result or Impact category data. Consequently, the extent to which information, based on ISO/TS 14048 and ISO 14044, is available for ecoinvent datasets varies depending on the type of dataset.

For *unit processes* the information that is needed according to ISO/TS 14048 should be available when combining the information implemented in datasets and in the reports. The aim with the reports is to provide a transparent description of the datasets. However, due to the fact that the information is distributed into two media, the information is not easily accessible and interpretable. The documentation of ecoinvent datasets as such in the EcoSpold format is not very detailed and lacks part of the information according to ISO/TS 14048. The major part of the documentation of datasets is available in separate reports, and to assess the applicability and quality of the datasets you would in general need to go to these reports for further information. However, the reports are a bit differently structured, and it may therefore not be easy to find the additional information about a datasets that may be needed in order to assess the quality of data in line with quality aspects according to ISO 14044.

The documentation of *cumulative LCI results and LCIA result*s lacks transparency, and does not follow requirements and recommendations for documentation of LCA

product systems according to ISO/TS 14048 and ISO 14044. The cumulative LCI results and LCIA results are generated automatically in the database by a calculation with a unit process as starting point, i.e. for each unit process both cumulative LCI results and LCIA results can be generated. There is, however, no information available about the result with regard to e.g. which processes that are included, system boundaries, etc. Some information from the interpretation of LCI results may be available in the separate reports, but it is generally quite scarce. Thus, it is difficult to assess the quality and use these types of datasets in LCA studies.

The available documentation for *Impact category datasets*, i.e. impact assessment data, is similar as for unit processes, i.e. a small part of the documentation has been implemented in the database in the EcoSpold format, whereas the major part of the documentation is available in a separate report. Thus, in the same way as for unit processes the information that is needed according to ISO 14044 may be available when combining the information in datasets and reports, but it is not in an easily accessible and interpretable form

Documentation in the EcoSpold format

The documentation of datasets in the EcoSpold format only includes a small amount of the information available for the datasets. The rest of the information is documented in separately available reports (see the section “Documentation in reports” below). According to the quality guidelines in Frischknecht et al. (2004), there are requirements on mandatory documentation, i.e. fields in the EcoSpold format that should be filled in. Nearly the entire EcoSpold format is mandatory for documentation of unit processes and product systems, except for three optional parts: Representativeness, Validations and Allocations. However, the completeness of the documentation differs depending on the type of dataset that is documented; i.e. depending on if it is a unit process, cumulative LCI result, LCIA result or impact category dataset.

With regard to ISO/TS 14048, the ecoinvent datasets generally lack information about choices of system boundaries, the purpose and specific applicability of the dataset, and the sources that have been used. This is mainly due to the fact that the EcoSpold format lacks structured fields for this information. However, depending on who is responsible for the documentation of the dataset, some of this information may be available in some part of the format for some datasets.

Unit processes and multi-output processes

In general, the documentation of unit processes and multi-output processes contains brief information about the datasets, and can be considered as a small excerpt of the full information that is available in the separate reports (see section “Documentation in reports” below). Multi-output processes are unit processes with more than one product. They are used to derive unit processes, i.e. processes with only reference function, through allocation. The allocation factors that are used to derive unit processes for each product from the multi-output process are documented in the multi-output process. This information is generally not available in the documentation of the allocated process, and consequently to get information about the allocation you need to go to the corresponding unallocated multi-output process.

Cumulative LCI results

The documentation of cumulative LCI results lacks transparency. They include no information of which processes that have been included in the calculation of the result, and how they are interlinked. The actual documentation of an LCI result dataset only includes the documentation of the unit process for which the LCI results have been calculated. The only way to distinguish the cumulative LCI result from the unit process is by the field "type" in Meta information. For unit processes this field contains the value "1" (corresponding to "Unit process"), whereas for cumulative LCI result the field contains the value "2" (corresponding to "System terminated").

The calculation of the result is performed automatically in the database. The datasets are interlinked by means of input and output flows, but the database does not present which processes that are included and how the processes are interlinked when performing a specific calculation. In practice, to get an idea of which processes that are included in a calculation, the user need to manually construct the LCI flowchart by following inputs flows backwards and connecting the corresponding processes. But even then, it may be difficult to determine that the resulting flowchart is the same as the actual flowchart used in the calculation.

The lack of transparency implies that the documentation does not follow the requirements and recommendations of documentation of product systems according to the ISO 14040 series of standards. Since it is difficult to know which processes that are included, it is not possible analyse the result to e.g. know which of the included processes that may have had a great influence on the results. If a data item is changed in one dataset, it could potentially change the LCI result of all other datasets. It is also not possible to e.g. choose not to include electricity production, or to make any other conscious modelling choices as to which process to include or exclude from the LCI result. This makes interpretation in line with requirements in the ISO 14040 series of standards difficult or impossible, and hampers use of the datasets in LCA studies.

Note: According to personal communication with Rolf Frischknecht, some software suppliers that resell the ecoinvent database with their software packages have constructed and made LCI flow charts available for the cumulative LCI results. The LCI results may therefore be more transparent when purchasing the database with such a software package.

LCIA results

Users of ecoinvent can also download data in the form of LCIA results, which is unit process data or cumulative LCI results including a calculation of impact assessment results for all impact assessment methods implemented in ecoinvent. Similarly as for the "cumulative LCI results", the actual documentation of an LCIA dataset only includes the documentation of the unit process for which the IA result has been calculated. IA results can be calculated for both unit processes and cumulative LCI results. The only way to distinguish an IA result from the unit process is by the fields "type" (see cumulative LCI results above) and "impactAssessmentResult", which contains the value "true" for IA results. For LCI data (unit processes and LCI results) the field contains the value "false".

Impact category data

Impact category datasets, i.e. impact assessment (IA) data, are documented in the EcoSpold format, using the same fields as for LCI datasets. The IA datasets are, however, not documented in line with ISO 14044, i.e. the concepts in impact assessment according ISO 14044, i.e. impact category and category indicators, characterisation and weighting can not be easily distinguished from the documentation. This may be due to the fact that the EcoSpold format is not originally designed for impact assessment data.

Depending on the impact assessment method, the data is provided as characterisation factors or impact factors (which includes both characterisation and weighting), but the documentation of the datasets does not clearly specify which type of factor that is provided in the dataset. This has to be interpreted from the separate report describing how the impact assessment methods have been implemented.

A review of the different implemented impact assessment methods also showed that the documentation of the datasets differs between different methods. The documentation in the EcoSpold format is thus not consistent between different methods in the database. This especially concern information that should be available according to ISO 14044, i.e. information about impact categories and category indicators, etc. All information about how the data has been adapted from the original method reports and implemented in the database is available in a separate report. Information about adaptations etc that has been made is, however, generally not implemented in the datasets in the EcoSpold format.

Documentation in reports

More detailed information about the datasets is available in 32 separate reports (see table 4). The reports are available on the CD-ROM that is provided when purchasing a license for the database and contain further information about unit process and multi-output process datasets, and the methodological choices that have been made in inventorying these datasets. The information include a description of the datasets (part of which is implemented in the documentation of datasets in the EcoSpold format), the sources used, some considerations regarding data quality and applicability of the datasets. With regard to cumulative LCI results or LCIA results, however, there is generally no further information available in the reports, except for some datasets where some information regarding interpretation of selected parts of the LCI result may be available. They do not include any flow charts or descriptions of included processes for the cumulative LCI results. Regarding impact assessment data, there is a separate report describing assumptions etc. that have been made when implementing the impact assessment methods in the database.

Table 4. Available reports describing the datasets in ecoinvent, and the number of pages in each report.

Themes	No	Citation	No of pages
Implementation of Life Cycle Impact Assessment Methods	3	(Frischknecht et al. 2004)	116
Life Cycle Inventories of ...			
Energy Systems: Results for Current Systems in Switzerland and other UCTE Countries	5	(Dones et al. 2004a)	182
Energiesysteme (energy systems, reports in German):	6	(Dones et al. 2004b)	
Zusammenfassung ZielMethodik	6-I-III		65
Erdöl (crude oil)	6-IV	(Jungbluth 2003a)	327
Erdgas (natural gas)	6-V	(Faist Emmenegger et al. 2003)	195
Kohle (coal)	6-VI	(Röder et al. 2004)	314
Kernenergie (nuclear energy)	6-VII	(Dones 2003)	405
Wasserkraft (hydro power)	6-VIII	(Bolliger & Bauer 2003)	100
Holzenergie (wood energy)	6-IX	(Bauer 2003)	134
Wärmepumpen (heat pumps)	6-X	(Heck 2004b)	39
Sonnenkollektor-Anlagen (solar collectors systems)	6-XI	(Jungbluth 2003c)	45
Photovoltaik (photovoltaics)	6-XII	(Jungbluth 2003b)	143
Windkraft (wind power)	6-XIII	(Burger & Bauer 2004)	83
Wärme-Kraft-Kopplung (combined heat and power plants)	6-XIV	(Heck 2004a)	145
Strommix und Stromnetz (electricity mixes and distribution)	6-XVI	(Frischknecht & Faist Emmenegger 2003)	102
Building Products	7	(Kellenberger et al. 2004)	714
Chemicals	8	(Althaus, Chudacoff et al. 2004)	927
Wood as Fuel and Construction Material	9	(Werner et al. 2003)	200
Metals	10	(Althaus, Blaser et al. 2004)	704
Packaging and Graphical Paper	11	(Hischier 2004)	
Introduction	11-I		17
Plastics	11-II		216
Paper and Board	11-III		335
Packaging Glass	11-VI		71
Detergents	12	(Zah & Hischier 2004)	112
Waste Treatment Services	13	(Doka 2003)	
General	13-I		111
Waste Incineration	13-II		97
Landfills	13-III		137
Wastewater Treatment	13-VI		59
Building Material Disposal	13-V		56
Transport Services	14	(Spielmann et al. 2004)	246
Agricultural Production Systems	15	(Nemecek et al. 2004)	289
Changes v 1.2	16	(Frischknecht et al. 2005)	98

All in all, there is a lot of further information available in these reports. However, the content of the different reports are not structured in the same way and some of the reports are only available in German (the reports describing energy and fuel systems). This may lead to the fact that it may be difficult to find the information that you are looking for in the reports, even if it may be available.

SPINE@CPM

The SPINE@CPM database contains more than 500 datasets. The data is implemented only as datasets documented in the SPINE format, in which all available information for the datasets is documented and published. There are no separate reports describing the datasets. The documentation of datasets does, however, contain references to the original sources that have been used in the compilation and documentation of the dataset, together with descriptions of how the references have been used. As mentioned in the section “Quality routines”, the quality requirements for SPINE@CPM are defined as detailed and structured documentation requirements, i.e. mandatory information that should be available for each dataset. The general and explicit purpose for these requirements is to make sufficient information available to allow for data quality assessments in line with ISO 14044.

The database contains different types of datasets with different scopes; both unit processes and LCI flow charts. For unit processes the scope ranges from individual unit operations, gate-to-gate processes, to complete cradle-to gate processes for a specific material. Some datasets are transparently documented as processes that consist of a combination of documented processes, i.e. as transparent LCI flow charts where each of the included datasets are documented separately.

The quality of the documentation of datasets varies, and there is a classification of the quality of documentation which provides an indication of the availability and sufficiency of documentation. In general, however, the documentation of datasets published in the database should include adequate or sufficient information to assess quality in line with the ISO 14044 standard. This of course, would depend on the detailed quality requirements for the specific study in which the dataset is to be used. Depending on specific requirements, there may be a need to go back to the original sources for the dataset for further information.

Impact assessment datasets are documented in the IA98 format, and contains transparent information in line with ISO 14044. In the documentation it is possible to clearly distinguish impact categories, category indicators, characterisation and weighting for the different impact assessment methods that are implemented. As for LCI-datasets, the documentation of IA datasets also includes references to the original sources that were used.

Conclusions

The ISO/TS 14048 data documentation format can be used as a tool for data quality management in LCA. The format can be used as a checklist of relevant information that is needed to be able to assess different data quality aspects in accordance with ISO 14044. As such, it can be used to assess the availability of relevant information and to identify if relevant information is missing in different data sources and databases.

By using the ISO/TS 14048 format as basis for the quality assessment of the ecoinvent and SPINE@CPM databases, different issues in regards to availability of information was identified.

The two databases have several similarities in regards to organization of the administration and have well established routines, but the character of quality requirements, documentation and data acquisition is different. The quality guidelines for ecoinvent are focused on modelling to harmonise the datasets in the database, whereas the quality requirements for SPINE@CPM are focused on structured documentation to facilitate interpretation of the different modeling choices that have been made in the compilation.

With regard to the data formats that are used in the databases, EcoSpold and SPINE, both formats lacks structured field for parts of the information according to ISO/TS 14048. The EcoSpold format, however, lacks more information than SPINE. This may be due to the fact that the EcoSpold format is intended as a data exchange format, and not as a data documentation format, which is the scope for both ISO/TS 14048 and SPINE.

The aim for both databases is to make transparent information available about the datasets, to allow for independent assessment of the quality and applicability of datasets. In the ecoinvent database, the information is distributed into two media (i.e. brief information in the EcoSpold format and further information in separate reports), whereas the SPINE@CPM is only available in one (i.e. all information in the SPINE format). The fact that two media is used in ecoinvent can hamper the accessibility of information.

The transparency of different types of datasets in the ecoinvent database also varies substantially; unit process, multi-output process and impact category datasets should be transparent when combining information in EcoSpold and reports. Cumulative LCI results and LCIA results, however, are not transparent since they lack information about the results, e.g. which processes that are included. In SPINE@CPM, transparent information should be available for all published datasets.

References

Arvidsson P. (Editor) et al (1997) *Krav på datakvalitet CPM:s databas 1997.* CPM Report 1997:1.

Arvidsson P., Carlson R., Pålsson A-C (1999) *An interpretation of the CPM use of SPINE in terms of the ISO 14041 standard.* CPM Report 1999:9.

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.

Carlson R., Pålsson A-C. (2002) *Documentation of environmental impact assessment, compatible with SPINE and ISO/TS 14048.* IMI-report 2002:1, Industrial Environmental Informatics (IMI), Chalmers University of Technology, Göteborg.

Carlson R., Steen B. (1998) *A Data Model for LCA Impact Assessment.* The 8th Annual Meeting of SETAC-Europe 14-18 April 1998, Bordeaux.

Erlandsson M., Pålsson A-C., Häggström S. (2006) *Specification of data conversion from EcoSpold to ISO/TS 14048, SPINE and IA98*. CPM Report 2006:1

Erixon M. (Editor) et al. (2000) *Facilitating Data Exchange between LCA Software involving the Data Documentation System SPINE*. CPM Report 2000:2.

Flemström K., Pålsson A-C. (2003) *Introduction and guide to LCA data documentation using the CPM documentation criteria and the ISO/TS 14048 data documentation format*. CPM report 2003:3

Frischknecht R., Althaus H.-J., Doka, G., Dones, R., Heck T., Hellweg, S., Hischier, R., Jungbluth N., Nemecek, T., Rebitzer, G., Spielmann, M. (2004) *Overview and Methodology. Final report ecoinvent 2000 No. 1*. Swiss Centre for Life Cycle Inventories, Duebendorf.

Hedemann J., König U. (2003) *Technical Documentation of the ecoinvent Database. Final report ecoinvent 2000 No. 4*. Institut für Umweltinformatik, Hamburg, DE, Swiss Centre for Life Cycle Inventories, Duebendorf, CH, for ecoinvent members only.

ISO 14044:2006 *Environmental management – Life cycle assessment – Requirements and guidelines*. International Standardisation Organisation

ISO/TS 14048:2002(E) *Environmental management – Life cycle assessment – Data documentation format,* International Standardisation Organisation

Pålsson A-C. (1999a) *Introduction and guide to LCA data documentation using the CPM data documentation criteria and the SPINE format*. CPM Report 1999:1.

Pålsson A-C. (1999b) *Review of LCI-data at SPINE@CPM*. CPM Report 1999:8.