

Integrating ERP and Environmental Information Systems – the Case of Life Cycle Inventories

Marcel Severith¹, Lorenz M. Hilty²

Abstract

This paper takes the life cycle inventory database *ecoinvent* as an example of the integration of ERP and environmental information systems. A reference model describes the ERP-integrated creation of life cycle assessments based on *ecoinvent* data. The main challenges to implement the reference model from the viewpoint of *ecoinvent* are identified and approaches for a solution presented. First, an outline for an ontology enhancing *ecoinvent* metadata is given to help bridge the so-called semantic gap. Second, the basic web services needed to attach to a service-oriented architecture are described.

1. Introduction

1.1 Motivation for environmental management

Respecting environmental limits and using energy and raw materials more efficiently can bring considerable cost savings. It can also help an enterprise gain competitive advantages especially with politicians and the public calling for more ecological products and services (Beucker 2000, 323-324).

However, ecological aspects are not always a priority of top management, although management may claim otherwise. In reality, often only compliance with environmental legislation is monitored instead of preventing environmental damages strategically (Junker 2010). Not even risk management takes any real notice yet of environmental information (Funk/Möller/Niemeyer 2007).

One reason for this disregard is inadequate IT support. So-called environmental management information systems (EMIS) are meant to support companies in matters of environmental protection (Hilty & Rautebrauch 1997; Hilty 2000). Unfortunately these systems are often isolated solutions, each solving sub-tasks only. If at all, they are coupled only loosely to other enterprise information systems (ERP), which are of central importance for modern businesses (Junker 2010; Teuteberg/Gómez 2010, 7-11).

Therefore, integrating means of environmental assessment into ERP systems can be regarded as an important step towards holistically incorporating environmental limits into operative and strategic decision processes.

1.2 Integrating environmental data into ERP – state of the art

The central challenge of integrating environmental assessment appears to be expanding the ERP master data. Normally ERP does not contain any environmental information at all. Furthermore data is often not

¹ Empa Material Science and Technology, Lerchenfeldstrasse 5, CH-9014 St. Gallen, marcel.severith@online.de

² Informatics and Sustainability Research (ISR), University of Zurich, Binzmuehlestr. 14, CH-8050 Zurich, and Empa Material Science and Technology, Lerchenfeldstrasse 5, CH-9014 St. Gallen, lorenz.hilty@empa.ch

as detailed as necessary for environmental assessment (Eul/Rey 2000, 214-222; Isenmann/Rautenstrauch 2007, 76-79).

Various studies have presented how environmental data could be integrated into ERP systems. Some are general models (Lang-Koetz/Heubach 2007), some concrete case studies on systems like infor:COM (Eul/Rey 2000) and SAP (Kuhrke/Feickert et al. 2008; Funk/Gómez et al. 2010).

However, none of these studies pay close attention to the source of environmental data. Often a vague explanation is given that data have been supplied by an EMIS, a web service or manual input. In the following, therefore, we describe what it looks like to acquire data from an environmental information system taking *ecoinvent* as an example.

1.3 Life Cycle Assessment with *ecoinvent*

Life Cycle Assessment (LCA) is a standardized method for estimating the environmental impacts of products and services. Therefore material and energy flows occurring along the whole life cycle are counted in relationship to the product output (Klöpffer/Grahl 2009, 1-4). LCA is not explicitly concerned with management structures or environmental plans (Krcmar/Dold et al. 2000, 15-16). However, an LCA can provide a prerequisite for environmental management, i.e. by identifying crucial factors for environmental improvements.

Within LCA there is often a lack of information regarding material and energy flows. While an enterprise may get specific data for some of its own activities, the environmental impacts of other own activities as well as impacts of upstream product chains have to be estimated with the help of generic background data (Klöpffer/Grahl 2009, 125-134).

ecoinvent is one of the world's leading databases for generic life cycle inventory data. With its comprehensive basis of over 4000 datasets (Frischknecht/Jungbluth et al. 2005; Klöpffer/Grahl 2009, 138-139; Hilty 2008, 137-138) *ecoinvent* can serve as a basis for LCA in cases where there is no specific data available.

1.4 Steps towards an integration of ERP and *ecoinvent*

The reference model depicted here describes the ERP-integrated creation of a simplified LCA with the help of *ecoinvent*. It is based on the model of Funk/Möller/Niemeyer (2009). The model assumes that the ERP is the leading corporate information system and thus also responsible for the process of LCA. Supporting systems like *ecoinvent* are integrated with the help of a service-oriented architecture (SOA).

Previous studies presented how environmentally relevant activities can be retrieved from the ERP master data; see for example Eul/Rey 2000; Isenmann/Rautenstrauch 2007; Kuhrke/Feickert et al. 2008; Funk/Niemeyer 2010. In addition, supplier-specific data can be used whenever it is available. Even downstream activities can be taken into account with the help of representative data and assumptions.

Thus we argue that the ERP can determine all relevant activities occurring throughout the life cycle of the given product. However, the environmental impacts of the activities remain unknown for the ERP. Therefore *ecoinvent* is used to gather the appropriate data and integrate it into the ERP.

For that the ERP activities have to be mapped to the activities represented in *ecoinvent* first. Thus far, it has not been possible to automatically collate data as part of a generic solution. Instead this task requires a lot of manual intervention and expert knowledge. The precise mapping of activities is hindered by different methodical assumptions and the lack of a common vocabulary. This problem is called the semantic gap. It has previously been identified as the biggest single challenge for an automated life cycle assessment by Funk/Möller/Niemeyer (2009).

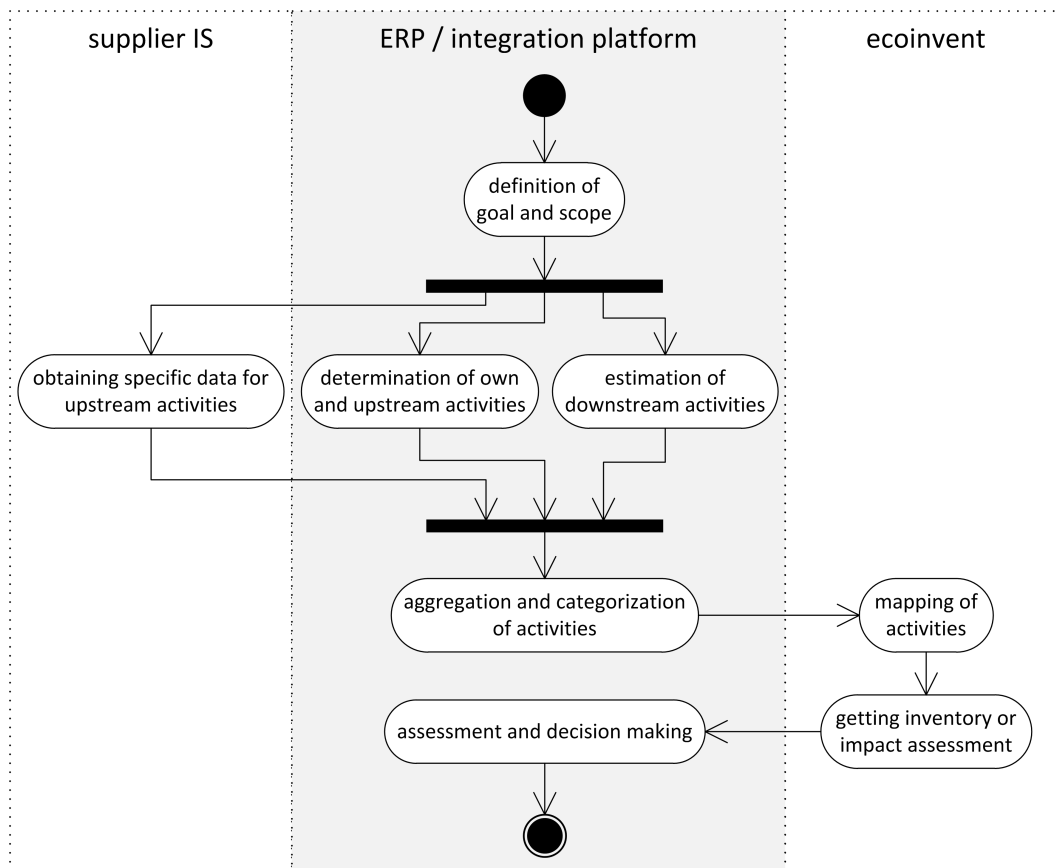


Figure 1
Activity diagram for creating an ERP-integrated LCA

Finally after mapping the activities, the environmental impacts can be retrieved from *ecoinvent*. While other models are typically rather strict with defining the form of representation, we advocate methodical openness. *ecoinvent* offers the aggregation levels of life cycle inventories as well as various methods of impact assessment. The form of representation varies with the needs of the subsequent analysis. For simplified LCA's the use of a simple method like the carbon footprint may be sufficient. For a more extensive, detailed analysis a full life cycle inventory may be necessary instead.

After that, the environmental impacts have to be analyzed. This again is seen as a task of the ERP system. The impacts of the individual activities have to be aggregated and visualized. For a detailed analysis individual aspects, such as components, can be accounted individually. The information gathered can then be used for decision making resulting in operational and strategic changes.

There are still many open challenges for both ERP and *ecoinvent* to actually implement the model. In the following we give some hints on how to solve the challenges from the perspective of *ecoinvent*.

2. Bridging the gap with the help of semantic data

As a system specialized in the representation of environmental data, *ecoinvent* should take an active part in bridging the semantic gap. One possible approach to bridge the gap is to semantically enhance the data of *ecoinvent*. The two following issues can be identified as the main challenges:

1. The activities in the ERP system and in *ecoinvent* are described in a different terminology.
2. The abstraction levels at which the activities are represented are different in each system. The environmental effects of an ERP activity may be described by multiple activity datasets in *ecoinvent*.

2.1 Existing metadata in *ecoinvent*

ecoinvent provides extensive metadata that is expected to enable proper application of the data. Most information is provided in the form of free text. According to Bischof/Bauer-Messmer (2008) it is easy for experts to input free text information. However, this form of representation hinders automatic processing as well as the understanding of users who lack the specialized terminology.

However, some metadata is semantically of particular interest already. A rudimentary taxonomy can be derived from the dataset structure of the current *ecoinvent v2*. The activities are categorized according to a proprietary scheme. It distinguishes among the thematic areas depicted in Figure 2.

1 agricultural means of production	15 heat pumps	29 plastics
2 agricultural production	16 hydro power	30 private consumption
3 biomass	17 insulation materials	31 solar collector systems
4 building components	18 lignite	32 textiles
5 chemicals	19 mechanical engineering	33 transport systems
6 construction materials	20 metals	34 ventilation
7 construction processes	21 mortar and plaster	35 washing agents
8 cooling	22 natural gas	36 waste management
9 district heating	23 nuclear power	37 water supply
10 electricity	24 oil	38 wind power
11 electronics	25 others	39 wood energy
12 food industry	26 paintings	40 wooden materials
13 glass	27 paper & cardboard	
14 hard coal	28 photovoltaic	

Figure 2
Top level of the current classification for *ecoinvent* activities

In *ecoinvent v3*, anticipated for release in mid-2012, it is intended to classify the activities on the basis of the ISIC standard (International Standard Industrial Classification). This standard defines a universal classification for economic activities. The restructuring makes sense from an organizational standpoint, as the Editorial Board responsible for quality control is also structured in conformity to the ISIC standard (*ecoinvent Centre n.d.*).

However, the contents of the individual activities can only be distinguished further by their name and their free text documentation. An explicit differentiation on the basis of the existing or upcoming classification is not possible.

Section	Divisions	Description
A	01–03	Agriculture, forestry and fishing
B	05–09	Mining and quarrying
C	10–33	Manufacturing
D	35	Electricity, gas, steam and air conditioning supply
E	36–39	Water supply; sewerage, waste management and remediation activities
F	41–43	Construction
G	45–47	Wholesale and retail trade; repair of motor vehicles and motorcycles
H	49–53	Transportation and storage
I	55–56	Accommodation and food service activities
J	58–63	Information and communication
K	64–66	Financial and insurance activities
L	68	Real estate activities
M	69–75	Professional, scientific and technical activities
N	77–82	Administrative and support service activities
O	84	Public administration and defence; compulsory social security
P	85	Education
Q	86–88	Human health and social work activities
R	90–93	Arts, entertainment and recreation
S	94–96	Other service activities
T	97–98	Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use
U	99	Activities of extraterritorial organizations and bodies

Figure 3
Top level of ISIC classification (United Nations 2008, 43)

2.2 Introducing a new classification of activities

In contrast to that, the semantics proposed here takes into account the various types of *ecoinvent* activities. Some of the conditions of the individual activities vary quite a lot, making an exact categorization of the activities extremely challenging. Nevertheless they can be subdivided into the following four main groups.

providing a product: most of the *ecoinvent* activities describe the providing of products (goods and services). For example, this type includes activities describing the extraction of raw materials, the production of components or the supply of electric energy. In some cases multiple activities describe providing same or similar products under different conditions and thus varying environmental impacts. In the case of raw materials, for instance, a distinction is made between primary raw materials and recycled secondary raw materials. Activities may already contain product losses, activities of wholesale and retail dealers as well as transport and disposal activities. In addition to the production of tangible goods providing a service falls into this category as well. Services can be recognized by the dimensions in which they are commercially measured. Whereas quantities of tangible goods are specified in pieces or physical mass, services are defined in a variety of other dimensions, such as distances, work or time.

processing a product: such activities describe the processing of products; especially in the area of raw materials. Normally such an activity contains only the environmental impacts of processing, but not the environmental impacts of providing the product under study.

disposal of a product: this type describes the environmental impact caused by disposing a product under specified conditions. This means in particular the type of disposal (incineration or landfill).

transport: transport, in deviation from the preceding methods, is usually not defined for specific products, but generically. Instead environmental impacts of transportation are defined as a function of transport means, distance and weight.

2.3 Introducing the concept of the product

All activities represented in *ecoinvent*, with the exception of transport activities, refer to the concept of the product. However, until now there has been no explicit mention of “product” in the database. Therefore let us provide an introduction in the following.

A thesaurus is used to define synonyms and relations between broader and narrower terms within a given domain. The use of an existing thesaurus to classify products would be an advantage. However Bischof/Bauer-Messmer (2008) have judged that services such as the thesaurus GEMET are unsatisfactory on account of its poor data quality. As for the German language, there is also the Semantic Network Service (SNS), which serves as a reference for systems of the Federal German Government (Rüther/Bandholtz 2008).

Many useful thesauri cover quite a narrow domain. In contrast *ecoinvent* claims to thoroughly cover the environmental impacts of providing goods and services as completely as possible. This requires semantic information across all economic areas.

Domain-independent data sources play an important role for the requisite connection of various data sources into a coherent network. One well-known example of that is the Project DBpedia, which provides structural data extracted from Wikipedia (Bizer/Becker 2010). Semantic processes can be improved with the large Wikipedia database. The individual Wikipedia articles have their own metastructure; the articles are linked logically with others and usually are even categorized. Furthermore, Wikipedia is extremely efficient in solving word forms having multiple meanings (Gillmeier/Hengartner/Pedrazzini 2010).

A first version of an extensive thesaurus can be made on the basis of categorizations in Wikipedia (Gillmeier/Hengartner/Pedrazzini 2010). However, it has to be expected that further manual intervention will be needed to satisfy specific requirements.

2.4 Introducing an ontology for *ecoinvent*

An ontology is a descriptive model representing the basic concepts, relations and constraints used to describe a given domain. It is shared by a group of people who want to exchange knowledge about the domain (Aßmann/Zschaler/Wagner 2006, 255-257). The ontology presented here is intended to serve as a semantic complement to the metadata existing in *ecoinvent*, bridging the semantic gap between *ecoinvent* and the ERP. The following aspects are transparent in particular:

1. *ecoinvent* activities are put into the predefined categories.
2. The product concept is introduced and made semantically useful with the aid of a thesaurus. Referring to a DBpedia entity accomplishes an unequivocal definition of a class of products.
3. Products and activities are brought into connection with each other.

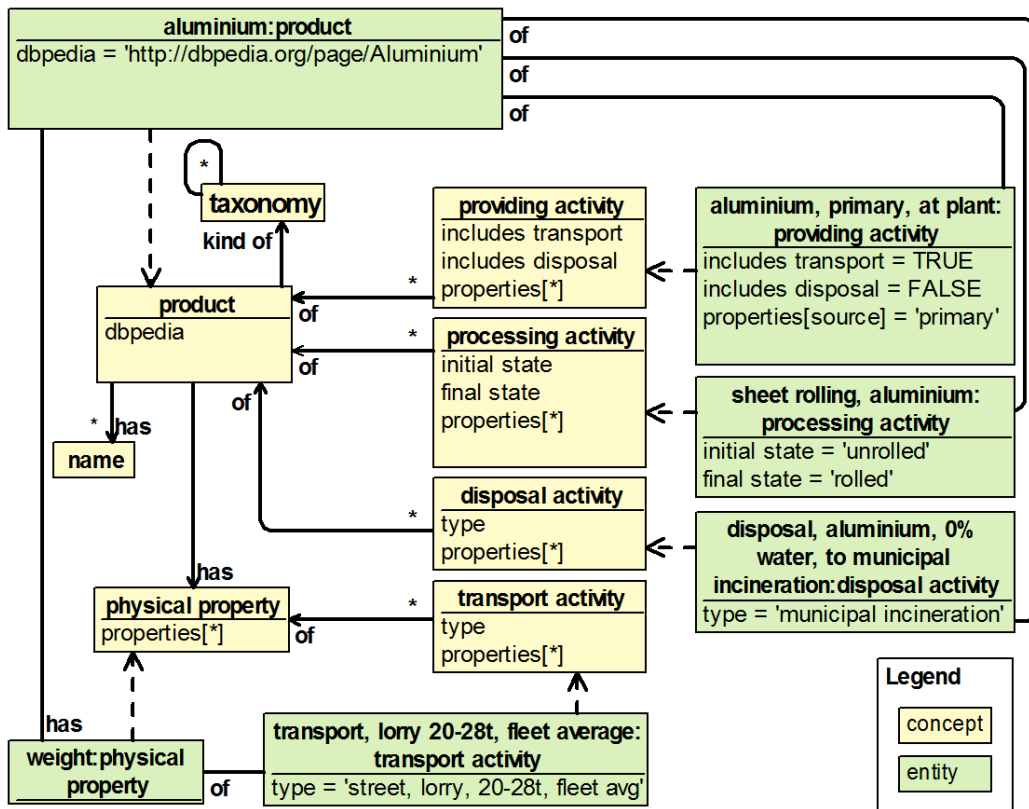


Figure 4
Approach of an ontology for semantic enhancement of *ecoinvent* data

The ontology makes it possible to take a product and to identify all corresponding *ecoinvent* activities for providing, processing and disposing it. In addition, suitable transport activities can be inferred from the physical properties of the product. In some cases multiple activities may exist per category that can be distinguished by their properties.

Under the assumption that the entity of the product and the type of activity can be identified in the ERP, mapping activities between ERP and *ecoinvent* becomes much easier. However, the following challenges have not yet been solved satisfactorily by the ontology:

1. The atomicity of products is still work in progress. We assume a certain abstraction level with our product concept. This means that similar products with different properties are collected under one product. “Aluminum” thus means a raw material independently of whether it was manufactured as primary raw material or secondary raw material.
2. For the unequivocal identification of activities their properties are to be defined exactly. These depend on both the activity category and the type of product. For example, in the case of raw materials, a distinction is made between primary or secondary raw materials, whereas agrarian products are classified according to cultivation methods. The difficulty exists here in the exact classification of products. As *ecoinvent* describes the “whole world” of tradable products, classification systems of specific disciplines soon reach their limits. Therefore designing a master system of classification is quite difficult.
3. The ontology describes only activities that are defined in *ecoinvent*. The ontology therefore naturally hits its limits whenever activities are lacking for specific products or product properties. For example,

there is no product defined such as an “aluminum capsule“. The ontology makes it easier to assemble “aluminum, primary, at plant” and “sheet rolling, aluminum” to an “aluminum capsule, of primary raw material”; but the links have to be made manually. The decision to include such an activity in *ecoinvent* and thus in the ontology may be justified by the existence of sufficient demand for such activities. Ultimately, though, the decision is made by the Editorial Board of *ecoinvent*.

3. Attaching to a service-oriented architecture (SOA)

Hansen/Neumann (2009, 664-665) speak of a trend toward modularization through service-oriented architecture (SOA) in ERP systems. A SOA offers prepackaged services via standardized interfaces, and thus improves linking of various applications. Therefore a SOA is also advisable for integrating *ecoinvent* in ERP systems. The required services can be identified with the help of the reference model described above.

3.1 Services to bridge the semantic gap

An important service for bridging the gap is getting a list of semantically related *ecoinvent* activities based on ERP data. This service accesses the predefined ontology. In addition to mapping the activity object, unit of measurement and geographic space also have to be determined. We argue, however, that these challenges are much easier to solve. As long as the automatic mapping is not sufficiently reliable, the user can be offered similar activities for manual verification in a semi-automatic process. On the long term an automatic correspondence can be set up by this service.

After manual verification, there may be a need to extend the ontology or even the *ecoinvent* database. In this way the future mapping of similar activities can be facilitated. Thus a service should also report the parameters of the mapping back to *ecoinvent*. One could think into the direction of “user generated content” and “collective intelligence” known from so-called Web 2.0 applications. Such an approach would be quite new for the review-based working method of the Editorial Board of *ecoinvent*. But a smooth integration of such new elements could help enhance the practical usefulness of *ecoinvent*.

3.2 Services for getting environmental data

At present *ecoinvent* provides its data in the EcoSpold format on the basis of XML. The XML file can be imported and post-processed by third-party software (Frischknecht/Jungbluth et al. 2005). Until now, it has still been necessary to select and download the files from the *ecoinvent* Website. This functionality can, however, be transformed into a service. This service can easily provide environmental data for individual activities in the following levels of detail:

1. cumulative results of life cycle inventory, a list of all environmental flows
2. cumulative results of life cycle assessment, based on various impact assessment methods

Additionally a service doing the impact assessment based on a life cycle inventory composed outside of *ecoinvent* may be useful. Thus the step of getting inventory data would not be solely dependent on *ecoinvent*. This would allow integrating other data sources as well.

For all these services the EcoSpold format is sufficient. In order to further the idea of a simplified LCA, however, an adaption of EcoSpold should be considered. Until now, EcoSpold offers only multiple methods of impact assessment at once. Given the fact that only one is usually needed, the desired method could be specified as a parameter. That would facilitate the data storage in ERP.

4. Conclusions and outlook

We presented a reference model for the ERP-integrated creation of simplified LCA. The model does not presuppose the existence of any dedicated environmental management information system (EMIS). Instead the life cycle inventory database *ecoinvent* provides data about the environmental impacts of corporate activities. The data can then be post-processed in the ERP.

This paper describes work required on the *ecoinvent* side to implement the reference model. That work consists mainly of bridging the semantic gap and providing web services to attach to a service-oriented architecture.

The web services needed for ERP integration have been described roughly. They serve both to map activities and to obtain environmentally relevant data. As for the services to obtain the environmentally relevant data, one can for the most part fall back on the implemented functionality and the internationally renowned data format EcoSpold.

For bridging the semantic gap we suggest to perform a semantic enrichment of *ecoinvent* data. First steps towards an ontology have been described. One of them is to introduce the concept of a product, another is to classify activities by their types and a third brings activities and products into connection with each other. The ontology we have presented makes it possible to see the broad outlines of a semantic search function. With semantic search the results can be improved in comparison with a purely text-based search. Due to a few still open questions, enjoying completely automatic mapping remains a distant perspective. However, one can assume that the ontology will at least facilitate the manual collation of activities.

Whether the ontology is feasible can only be verified using case studies. We expect that adaptations and expansions derived from practical experience will be necessary. For that, it appears promising to derive general requirements for semantics from different business sectors. Investigating multiple sectors will help to draw domain-independent conclusions.

Bibliography

- Aßmann, U., Zschaler, S., Wagner, G. (2006): Ontologies, Meta-models, and the Model-Driven Paradigm, in: C. Calero et al. (eds.): *Ontologies for Software Engineering and Software Technology*, Berlin Heidelberg, pp. 249-273
- Beucker, S. (2000): Umweltinformationen - entscheidender Faktor für den Unternehmenserfolg, in: H.J. Bullinger et al. (eds.): *Auftragsabwicklung optimieren nach Umwelt- und Kostenzielen*, Berlin Heidelberg, pp. 323-331
- Bischof, S., Bauer-Messmer, B. (2008): Semantic Enhancement of Environmental Metadata, in: A. Möller et al. (eds.): *Environmental Informatics and Industrial Ecology. 22th International Conference on Informatics for Environmental Protection. EnviroInfo 2008*, Aachen, pp. 123-131
- Bizer, C., Becker, C. (2010): Semantische Mashups auf Basis des Linked Data Web, in: U. Hengartner, A. Meier (eds.): *Web 3.0 & Semantic Web*, HMD, Heft 271, Heidelberg, pp. 59-69
- ecoinvent Centre (n.d.): Editorial Board, <http://www.ecoinvent.org/organisation/editorial-board/>, last access: 22. July 2011
- Eul, D., Rey, U. (2000): Einführung von umweltorientierten Funktionalitäten in ERP-Systemen, in: H.J. Bullinger et al. (eds.): *Auftragsabwicklung optimieren nach Umwelt- und Kostenzielen*, Berlin Heidelberg, pp. 205-225
- Frischknecht R., Jungbluth N., Althaus H.J., Doka G., Dones R., Heck T., Hellweg S., Hirschier R., Nemecek T., Rebitzer G., Spielmann M. (2005) The ecoinvent database: Overview and Methodological Framework, *International Journal of life cycle assessment*, Vol. 10, No. 1, pp. 3-9

- Funk, B., Gómez, J.M., Niemeyer, P., Teuteberg, F. (2010): Geschäftsprozessintegration mit SAP. Fallstudien zur Steuerung von Wertschöpfungsprozessen entlang der Supply Chain, Berlin Heidelberg
- Funk, B., Möller, A., Niemeyer, P. (2009): A reference architecture for the integration of EMIS and ERP-systems, in: *Lecture Notes in Informatics*, 154, pp. 3393-3401
- Funk, B., Möller, A., Niemeyer, P. (2007): Integration of Risk-Oriented Environmental Management Information Systems and Resource Planning Systems. In: O. Hryniewicz et al. (eds.): *Environmental informatics and systems research*, Band 1, Aachen, pp. 545-552
- Funk, B., Niemeyer, P. (2010): Abbildung von Umweltwirkungen in betrieblichen Informationssystemen, in: J.M. Gómez et al. (eds.): *Green Computing & Sustainability*, HMD, Heft 274, Heidelberg, pp. 37-46
- Gillmeier, S., Hengartner, U., Pedrazzini, S. (2010): Wie man mit der Wikipedia semantische Verfahren verbessern kann, in: U. Hengartner, A. Meier (eds.): *Web 3.0 & Semantic Web*, HMD, Heft 271, Heidelberg, pp. 70-80
- Hansen, H.R., Neumann, G. (2009): *Wirtschaftsinformatik 1. Grundlagen und Anwendungen*, Stuttgart
- Hilty, L.M. (2008): *Information technology and sustainability. Essays on the relationship between ICT and sustainable development*, Norderstedt
- Hilty, L. M. (2000): Umweltinformationen strategisch und betriebsübergreifend nutzen, in: A.B. Cremers, K. Greve (eds.): *Umweltinformation für Planung, Politik und Öffentlichkeit*, Umweltinformatik aktuell, Band 26, Marburg, pp. 716-722
- Hilty, L. M., Rautenstrauch, C. (1997): Environmental Information Systems for Production and Recycling, in: D. Swayne et al. (eds.): *2nd International Symposium on Environmental Software Systems (ISESS)*, New York, pp. 21-29
- Isenmann, R., Rautenstrauch, C. (2007): Horizontale und vertikale Integration Betrieblicher Umweltinformationssysteme (BUIS) in Betriebswirtschaftliche Anwendungsszenarien, in: *uwf*, Vol. 15, No. 2, pp. 75-81
- Junker, H. (2010): Die Beliebbarkeit betrieblicher Umweltinformationssysteme - eine Polemik, in: K. Greve, A.B. Cremers. (eds.): *Integration of Environmental Information in Europe. EnviroInfo 2010*, Köln Bonn, Aachen, pp. 232-247
- Klöpper, W., Grahl, B. (2009): *Ökobilanz (LCA). Ein Leitfaden für Ausbildung und Beruf*, Weinheim
- Krcmar, H., Dold, G., Fischer, H., Strobel, M., Seifert, E.K. (2000): *Informationssysteme für das Umweltmanagement. Das Referenzmodell ECO-Integral*, München Wien
- Kuhrke, B., Feickert, S., Stöcker, P., Melk, K., Clesle, F.D., Bachmann, E. (2008): Life Cycle Design auf Basis von Standardsoftwaresystemen, in: E. Abele et al. (eds.): *EcoDesign. Von der Theorie in die Praxis*, Berlin Heidelberg, pp. 149-184
- Lang-Koetz, C., Heubach, D. (2007): *Umweltcontrolling umsetzen - Erstellung von Kennzahlen für Stoff- und Energieströme und deren Integration in die betriebliche IT*, Stuttgart
- Rüther, M., Bandholtz, T. (2008): Semantic Network Service (SNS) Update and Future Vision, in: A. Möller et al. (eds.): *Environmental Informatics and Industrial Ecology. 22th International Conference on Informatics for Environmental Protection. EnviroInfo 2008*, Aachen, pp. 401-408
- Teuteberg, F., Gómez, J.M. (2010): Green Computing & Sustainability. Status quo und Herausforderungen für betriebliche Umweltinformationssysteme der nächsten Generation, in: J.M. Gómez et al. (eds.): *Green Computing & Sustainability*, HMD, Heft 274, Heidelberg, pp. 6-17
- United Nations (eds.) (2008): *International Standard Industrial Classification of All Economic Activities, Revision 4*, Statistical papers, New York