

ecoinvent 2000 - Documentation EcoSpold

Document version: 1.3
Authors: J. Hedemann, ifu
I. Meinshausen, ifu
Translation: R. Frischknecht, ecoinvent centre
Release date: 26. October 2008
File name: Documentation EcoSpold 1.3.English.doc

Contents

1	Introduction	2
2	References	2
3	General Information	2
4	Structure and Validation	3
4.1	Principle structure	3
4.2	Limits of XML schemas and their overcoming.....	3
4.3	Categories.....	3
5	Processes, elementary flows and impact categories	3
6	Files	4
7	Installation Microsoft XML Parser	6
8	Software components	6
8.1	Installation.....	6
8.2	Position of the EcoSpoldAccess Toolbar	7
8.3	EcoSpoldAccess Options	7
8.4	Export of Excel data into EcoSpold documents	8
8.4.1	Excel2EcoSpold	8
8.4.2	Excel2EcoSpold Impact	9
8.5	Import of EcoSpold documents	10
8.6	EcoSpold Access.....	10
8.6.1	Primary/Foreign key violations	12
8.6.2	Schema documentation.....	12

1 Introduction

The EcoSpold format defines a data exchange format for process data in XML. The conversion of EcoSpold documents from Excel-Files, the validation of EcoSpold documents, their display in HTML and their import back to Excel is enabled by EcoSpold. The installation includes further detailed documentation and some fictive examples.

After the successful installation all XML documents defining the EcoSpold format are in the main EcoSpold folder. With the help of these documents, EcoSpold documents are validated in terms of structure and content of the EcoSpold format. Chapters 4 to 6 describe the structure of the EcoSpold format and the files used for it.

The EcoSpoldAccess component for Microsoft Excel supports the following functions (only main functions are listed):

- Display of EcoSpold documents in HTML,
- Conversion of Excel data to EcoSpold documents and vice versa,
- Validation of EcoSpold documents,
- Display of EcoSpold schema documentation which include information about the content of the individual EcoSpold data fields.

These functions of the EcoSpold Access component are extensively discussed in chapter 8, where users find the important information about how to use the EcoSpold Access component.

2 References

This document is completed by the following documents:

- EcoSpold documents generated from the schemas (EcoSpold01Dataset.html, EcoSpold01ElementaryDataset.html and EcoSpold01ImpactDataset.html). These documents describe the content of the EcoSpold data fields.
- The EcoSpold Public License (ESPL), which explains the terms of use for the EcoSpold components (EcoSpold Public License v1.0.pdf).
- <http://www.w3.org/XML>
Official Website for XML (eXtensible Markup Language)
- <http://www.w3.org/XML/Schema>
Official Website for XML Schema
- <http://www.w3.org/Style/XSL>
Official Website for XSL (eXtensible Stylesheet Language)
- <http://www.w3.org/TR/xslt>
Official Website for XSLT (XSL Transformation)
- <http://www.microsoft.com/xml>
Website for the Microsoft XML Parser

3 General Information

The EcoSpold format is based on XML documents. XML documents can be validated with the help of rules defined with XML schemas and XLS transformations.

On May 2, 2001, XML Schema did reach the Status of a “W3C Recommendation“ (cf. <http://www.w3.org/XML/Schema>). They are on the way to a de facto standard for communication in the Internet (similar to HTML, CSS).

Besides their validation function, EcoSpold schemas are used for documentation purposes as well. The schema allows to read the principle structure of an EcoSpold document. In addition to that, all elements within a schema can be completed with documentations and comments. An XLS file has been developed to edit this documentation as a HTML file.

The assessment of existing XML specifications lead to the conclusion that no one was suitable to build upon. The EML of Humboldt University Berlin was not yet far developed and did focus on a completely different area. However, other specifications have been used to come as close as possible to "best practice".

4 Structure and Validation

4.1 Principle structure

Within an EcoSpold document further information and any number of EcoSpold datasets can be stored. Relevant is the complex datatype Tdataset, which describes one complete EcoSpold dataset. Elements of this datatype (dataset) may occur several times within one single document. During the import, each occurrence of dataset within the XML document is identified and each occurrence of dataset is imported as one single EcoSpold dataset.

4.2 Limits of XML schemas and their overcoming

Not all requirements of the EcoSpold format can be validated with the help of schemas. XML schemas are of no help in the following situations:

- A value range is only defined as a recommendation, not as a strict requirement. If the value is outside the recommended range, a warning is issued. Examples are data field No 662 "location" and data field No 301 "companyCode".
- A value range covers two or more data fields. Hence, valid value combinations should be required. However, in XML lists of valid values can only be defined for one single datafield. Examples: data fields No 495/496 category/subCategory.
- The dependence of admitted values on the content of other datafields can hardly be described. For that purpose, the xsd:choice element in the XML schema is foreseen, which allows for alternative validation requirements. An example of an application of xsd:choice can be found within the exchanges, data fields No 3503 (inputGroup) and No 3504 (outputGroup). An example where we did not apply xsd:choice is data field No 201 (dataset type). Here three different schemas have been developed.

The first two problems are solved with the help of XSLT. XSLT uses a language (XLS) to transform XML documents. It can be used to convert an XML document into an HTML document. XSL is also suited for semantic validations. If errors are found they are written in an XML document, which is the result of an XSL transformation.

4.3 Categories

In ecoinvent database, categories (and subcategories) are used to structure its content. Examples are "construction wood" or "electricity" on the level of categories, "extraction" or "supply mix" on the level of subcategories. Usually categories are project specific and therefore not an integral element of the ecospold format. A list of the ecoinvent 2000 categories and subcategories are stored in an XML document, that can be changed by the user.

5 Processes, elementary flows and impact categories

Depending on the dataset type (defined by data field No 201) very different demands are made on the existence of the various areas. It is therefore difficult to develop one single schema. If the schema defines an area (see Table 1 and EcoSpold01Dataset.html) as "required", information must be provided. If an area is only defined as "optional", because the area makes no sense for elementary flows and is therefore not required, validation becomes very weak.

That is why three separate schemas have been developed, one for processes, elementary flows and for impact categories. The areas are described in Table 1.

Area	0=System non-terminated, 1=Unit process, 5=MultiOutput	2=System terminated	3=Elementary Flow	4=Impact Category
DataSetInformation	always	always	always	always
DataEntryBy	always	always	possible	always
ReferenceFunction	always	always	always	always
TimePeriod	always	always	never	possible
Geography	always	always	never	never
Technology	possible	possible	never	never
Representativeness	possible	possible	never	never
DataGeneratorAndPublication	always	always	never	always
Sources	always	always	never	always
Allocations	possible	never	never	never
Exchanges	always	always	never	always
Validations	possible	possible	possible	possible
Persons	always	always	possible	always

Table 1: Required-Matrix for EcoSpold Areas

The area "Validations" is required for ecoinvent, but its application by other users should not be forced.

6 Files

The following files are currently part of the EcoSpold format:

Schema-File	Content
EcoSpold01Dataset.XSD	Contains the schema for a complete dataset of a process. Uses the other XSD files.
EcoSpold01ElementaryDataset.XSD	Contains the schema for a complete dataset of an elementary flow. Uses the other XSD files.
EcoSpold01ImpactDataset.XSD	Contains the schema for a complete dataset of an impact category. Uses the other XSD files.
EcoSpold01MetaInformation.XSD	Contains the schema for the metainformation of a process.
EcoSpold01FlowData.XSD	Contains the schema for the relations of a process.
EcoSpold01DataTypes.XSD	Defines the basic data types, including CountryCodes, LanguageCodes and Persons.
EcoinventCompanyCodes.XSD	Schema for the list of the valid Companies.
EcoinventRegionalCodes.XSD	Schema for the list of the valid Regional Codes.
EcoinventUnits.XSD	Schema for the list of the valid Units.
EcoinventCategories.XSD	Schema for the list of the valid Categories.

XML data files	Content
Categories.XML	List of valid categories, including information: Category, SubCategory, LocalCategory, LocalSubCategory.
CompanyCodes.XML	List of valid company codes, including information: company abbreviation (CompanyCode), country (CountryCode).
RegionalCodes.XML	List of valid regional codes completing the list of country

	codes used in data fields "Location" and including information: abbreviation, name.
Units.XML	List of valid units.
XSL files for display in HTML	
	Content
EcoSpold01.xsl	General stylesheet for the display of EcoSpold documents in HTML format. It includes a table of content at the beginning. Uses the XSL files described below.
EcoSpold01Data.xsl	Display of data of an EcoSpold (process) document.
EcoSpold01ImpactData.xsl	Display of data of an EcoSpold Impact Category document.
EcoSpold01ElementaryData.xsl	Display of data of an EcoSpold Elementary Flow document.
XSL Files for semantic validation	
	Content
EcoSpold01SemanticValidator.XSL	XSL file, that makes the semantic validation not covered by the schemas. Uses the other *SV.XSL files.
EcoSpold01SV.XSL	Validation of an EcoSpold (process) dataset.
EcoSpold01ImpactSV.XSL	Validation of an EcoSpold Impact Category dataset.
EcoSpold01ElementarySV.XSL	Validation of an EcoSpold Elementary Flow dataset.
XSL files for schema documentation	
	Content
EcoSpold01SchemaDocumentation.XSL	XSL Transformation, applied on the EcoSpold01.XSD file, to create the documentation.
EcoSpold01SchemaDocumentationData.XSL	Creates the data part of the documentation.
EcoSpold01SchemaDocumentationTOC.XSL	Creates the navigation part of the documentation.
EcoSpold01SchemaDocumentationSimple.XSL	XSL Transformation, applied on the EcoSpold01.XSD file, to create the print version of the documentation.
Example files (both *.xls and *.xml)	
	Content
EcoSpoldExample_HeatOil_v1.0 (suffix 'u': incl. uncertainty data)	The example contains fictive LCI data for the supply of useful heat with a light fuel oil boiler. The product system consists of crude oil extraction, long distance transportation, oil refining, regional distribution, and heat production. Additionally, steel production and grid electricity supply are modelled as background data. The example has been developed for and first been published in Schmidt & Schorb 1995.
EcoSpoldExample_RoundWood_v1.0 (suffix 'u': incl. uncertainty data)	The round wood example models forestry products. It includes fictive data for the construction of forest roads, cutting and transportation of trees as well as for the supply of round wood and residual wood at the forest road. It contains a multioutput process (supply of round and residual wood), and processes for the production and for the operation of a power saw.
EcoSpoldExample_Impact_v1.0	The example contains fictive characterisation and damage factors that are used for the impact assessment

in the ecoinvent demo data. The impact categories take pattern from the eco-indicator 99 method but the figures are invented.

Further files	Content
EcoSpold01.spp	Project file for the software XML spy.
ImportTemplate.xls	MS Excel format pattern for the import of EcoSpold (Process) documents in Excel.
ImportTemplateImpact.xls	MS Excel format pattern for the import of EcoSpold Impact Category documents in Excel.
ImportTemplateElementary.xls	MS Excel format pattern for the import of EcoSpold Elementary Flow documents in Excel.
EcoSpoldAccess.xla	MS Excel Add-In for an easy access to the EcoSpold functionalities (cf. chapter 8)
EcoSpoldAccess.dll	Dynamic Link Library for the EcoSpold Access component.
AddInCheck.exe	Routine for the registration of the EcoSpold Access component at MS Excel. Not suited for an isolated use outside the EcoSpold installation.

7 Installation Microsoft XML Parser

The installation of the Microsoft XML Parser 6.0 is required for the use of the EcoSpold Access component. It is included in Windows Vista and Windows XP SP3. A separate Setup is provided by Microsoft free of charge.

8 Software components

The EcoSpold Access component enables the display of EcoSpold documents in HTML format (MS Internet Explorer v5.5 or later versions required), the validation of EcoSpold documents, the creation and display of the EcoSpold schema documentations (MS Internet Explorer v5.5 or later versions required), as well as the import and export of EcoSpold documents to and from MS Excel. The installation of Microsoft Excel 2003 is required to use the EcoSpold Access component.

8.1 Installation

The installation of the EcoSpold Access component is made with a setup program that copies the required files, makes necessary settings and registers the toolbar at MS Excel. If the installation of the EcoSpold Access toolbar(s) fails, a manual activation is recommended. The installation program verifies whether earlier versions are installed and offers their deinstallation.

1. First make sure that the toolbar(s) of earlier versions have been removed. Choose the option "adjust ..." in the menu that pops up by a right click with the mouse on any toolbar. Be sure that the toolbar(s) Excel2EcoSpold and Excel2EcoSpoldecoinvent are not listed. If they are listed, remove them with the delete function.
2. Close MS Excel and remove all files with the name EcoSpoldAccess.xla from your computer.
3. Reinstall EcoSpold.
4. If the toolbar(s) do not show up after installation, the EcoSpold Access component needs to be activated manually via the Add-In manager (Extras | Add-In manager). If the EcoSpold Access component has already been activated in the Add-In manager, it should be deactivated and the Add-In manager shall be closed. After a reactivation of the EcoSpold Access component in the Add-In manager, the EcoSpold toolbar(s) are displayed in the bottom frame of Excel.

8.2 Position of the EcoSpoldAccess Toolbar

After a successful installation the EcoSpoldAccess toolbar(s) appear at the bottom frame of MS Excel (see Figure 1). Closing MS Excel, visibility and position of the toolbar(s) are saved in the registry and will be reread at the startup of MS Excel.

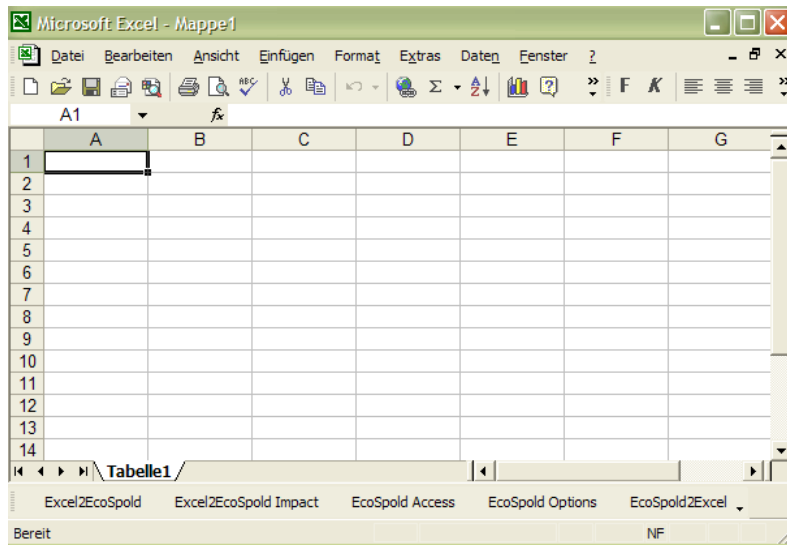


Figure 1: EcoSpoldAccess Toolbar (external version) after a successful installation.

The toolbar buttons of the external toolbar provides export- and comfort functionalities for all users. Their functions are described in sections 8.3 to 8.6.

The (additional) toolbar of the internal version provides functionalities to create root data for new versions of ecoinvent data and to manage names lists for processes, elementary flows, impact categories and database categories and subcategories.

8.3 EcoSpoldAccess Options

The EcoSpold Options button opens the Options dialogue where important settings for the EcoSpold Access component and the export functions can be made.

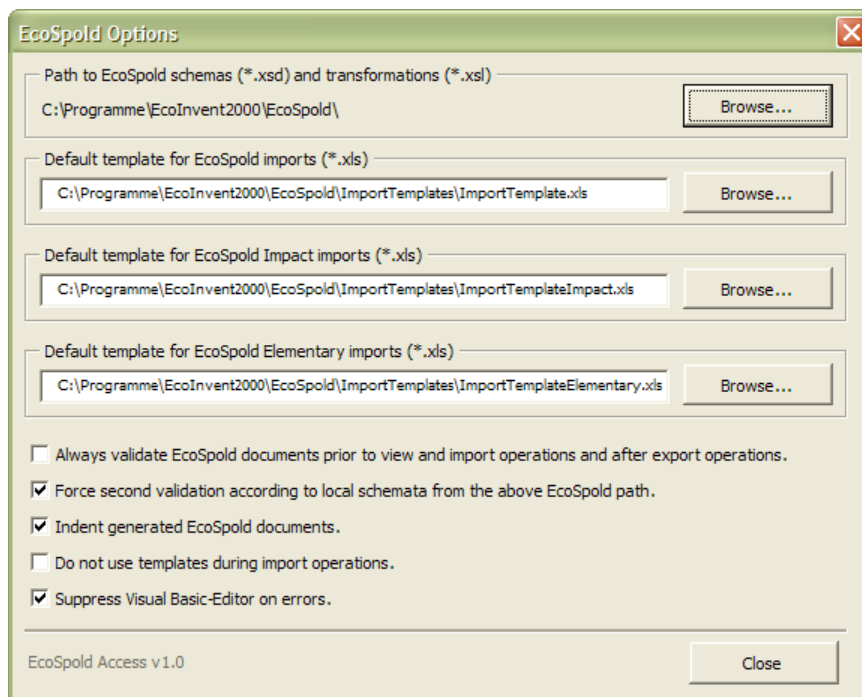


Figure 2: Options dialogue of the EcoSpold Access components

The value of the „Always validate...“ checkbox defines whether or not EcoSpold documents are validated before display via the EcoSpold Access component. This option is also valid before import and after export of EcoSpold documents to and from MS Excel tables.

The value of the „Force second validation...“ checkbox is only of interest for a validation of EcoSpold documents produced by third parties. Each EcoSpold document refers - via the xsi:schemaLocation attribute - to the location from which the XML parser shall load the XML schema for validation. Because the directories' structure of the local computer may differ from the one of the computer that produced the EcoSpold document, the reference may not show to the required XML schema on the local computer. In that case, a second validation is made based on the local XML schemas. The second validation is only made if the first validation fails because the value of the xsi:schemaLocation attribute does not refer to the local schemas.

The value of the „Indent generated EcoSpold documents“ checkbox defines whether or not all EcoSpold documents are saved with indents. Documents saved with indents are easier to view with the help of a simple text editor. Furthermore, indented EcoSpold documents can make use of meaningful line numbers for error reporting.

The option „Do not use templates...“ enables to obviate the use of the three templates (format patterns) for the import to MS Excel. Instead of these templates a predefined table is created and filled with import data.

The option „Suppress Visual Basic-Editor...“ enables to suppress the display of the Visual Basic Editor in cases of errors in the source code of EcoSpold Access component.

The setting „Significant digits for truncation of exchange values“ defines the number of significant digits used when converting Excel data into the EcoSpold format. This number of significant digits is valid for all numbers and percent values read from the X-Exchange table.

8.4 Export of Excel data into EcoSpold documents

All export functions require a corresponding open Excel workbook. If required worksheets are missing in the respective workbook, the user will be alerted and the export function aborted.

8.4.1 Excel2EcoSpold

The Excel2EcoSpold Export requires a workbook with the following sheets: X-Exchange, X-Process, X-Source and X-Person.

ID	Input Group	Output Group	Name	Location	Category	SubCategory	Infrastructure Process	Unit
401								
662								
493								
403								
5			round wood, hardwood, under bark, u=70%, at forest road	RER			0	m3
5			residual wood, hardwood, under bark, u=80%, at forest road	RER			0	m3
5			power sawing, with catalytic converter	RER			0	h
5			power saw, with catalytic converter	RER			1	unit
5			construction, forest road	RER			1	m
5			petrol, two-stroke blend, at regional storage	CH			0	kg
5			diesel, burned in building machine	GLO			0	MJ
4			Occupation, forest, intensive, normal		resource	land		m2a
4			Carbon dioxide, in air		resource	in air		kg
4			Methane, fossil		air	low population den:		kg

For each column No 3707 (name of the reference function) one dataset is created in the EcoSpold format, whose reference function has the value of the cell one line below the 3707 cell. The exchanges are exported to the EcoSpold document starting from line 7 until an empty line in column 6 (or F, Name). Exchanges are listed in the EcoSpold dataset if the corresponding cells in the 3707 columns do not contain one of the following values: “”, “0”, “-” or “”.

Uncertainty values, if available, need always be listed directly after the 3707 column. In the second line, the EcoSpoldIDs of the uncertainty data fields are listed and in the lines for the exchanges the uncertainty values are entered. The sequence of the uncertainty data fields is 3708, 3709, 3795, 3796, 3797 (UncertaintyType, StandardDeviation95, MinValue, MaxValue, MostLikelyValue).

For each coproduct of a multioutput process, a column with the EcoSpoldID 2404 is required placed after the 3707 column and eventually available uncertainty information of the multioutput process. In the 2404 columns allocation factors of the coproducts are listed. The output group of the coproducts listed in the exchanges is '2' (see also schema description).

After the creation of the EcoSpold document the user has the possibility to define the file name and to select the folder in which the document shall be saved. After the export a dialogue box reports the status of the export.

8.4.2 Excel2EcoSpold Impact

The Excel2EcoSpold Impact Export requires a workbook with the following worksheets (see also Figure 4): X-ImpactFactor, X-Process, X-Source und X-Person.

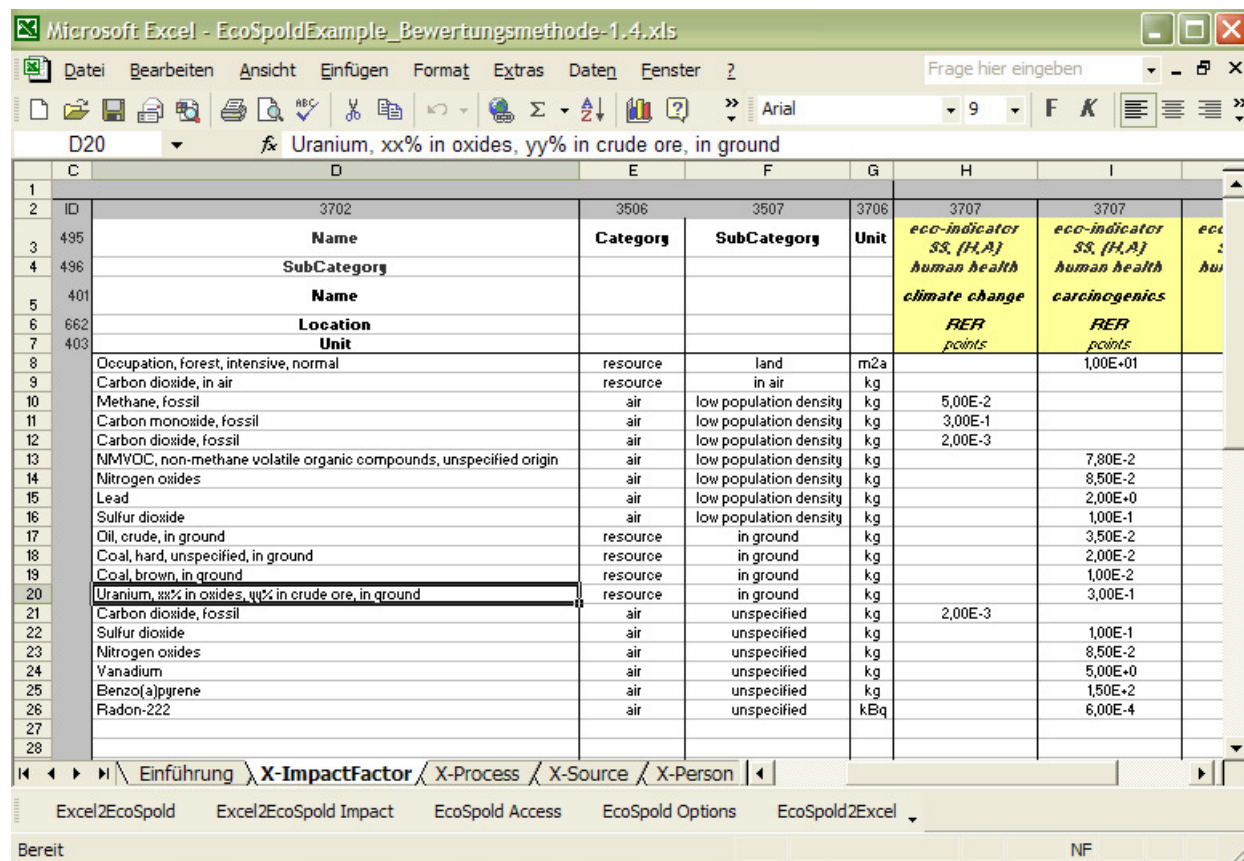


Figure 4: Excerpt of a correct Excel workbook for Excel2EcoSpold Impact

Similar to the Excel2EcoSpold export a dataset is created for every 3707 column. The exchanges (rather called characterisation, damage or weighting factors) are exported to the EcoSpold document starting from line 8 until an empty cell is reached in column 4 (or D). Exchanges are listed in the EcoSpold dataset if the corresponding cells in the 3707 columns do not contain one of the following values: “”, “0”, “-” or “”. No uncertainty information is admitted.

After the creation of the EcoSpold document the user has the possibility to define the file name and to select the folder in which the document shall be saved. After the export a dialogue box reports the status of the export.

8.5 Import of EcoSpold documents

The EcoSpold2Excel button imports existing EcoSpold documents into MS Excel. The type of the EcoSpold document (process, elementary flow, impact category) is recognised automatically and the Excel template is loaded according to the settings in the EcoSpold Options dialogue. The import function supports the three name spaces <http://www.Ecoinvent.org/EcoSpold01>, <http://www.Ecoinvent.org/EcoSpold01Elementary> and <http://www.Ecoinvent.org/EcoSpold01Impact>. If an invalid XML-document is selected for import, it may be displayed in the Internet Explorer.

After a successful import the Excel file can be saved with the help of a file saving dialogue under any name. If the saving procedure is aborted the created Excel file needs to be closed before a next import is made. The question in the Excel-dialogue whether or not to save the changes in the Excel file needs to be answered in the negative in any case because otherwise the used format pattern will be overwritten.

8.6 EcoSpold Access

EcoSpold documents can be displayed and validated with the help of the EcoSpoldAccess component (see Figure 5). Additionally, EcoSpold Schema documentations can be displayed with this component.

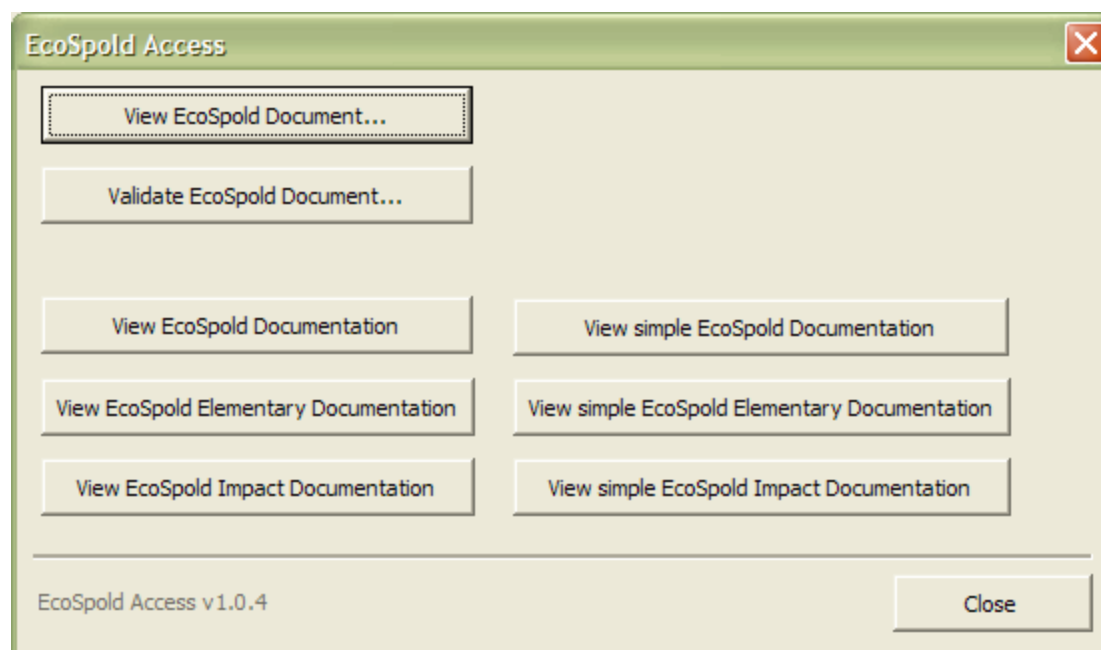


Figure 5: EcoSpold Access dialogue

All buttons with three full stops in the dialogue window shown in Figure 5 open another dialogue to choose a file (see Figure 6).

Before conversion to HTML the default name space of the selected XML document is validated. The conversion is aborted if the name space does neither correspond to „<http://www.Ecoinvent.org/EcoSpold01>“, nor to „<http://www.Ecoinvent.org/EcoSpold01Elementary>“ nor to „<http://www.Ecoinvent.org/EcoSpold01Impact>“. The file name under which the created HTML code of the EcoSpold document shall be saved can be entered in a dialogue window. The name of the EcoSpold document is proposed with the ending HTM. After the conversion via the EcoSpold Access component the HTML files can be displayed with the Internet Explorer and without the EcoSpold Access component.

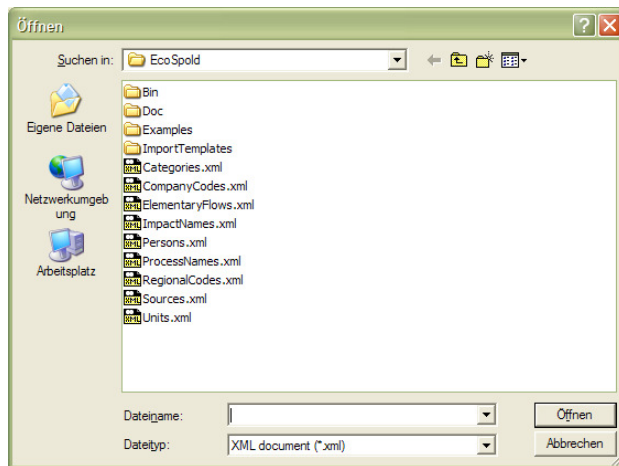


Figure 6: Dialogue "Open file"

The button „Validate EcoSpold Document...“ validates whether the EcoSpold document is well-shaped, i.e., whether it is a correct XML document. Afterwards the document is validated in view of the syntactical requirements of the EcoSpold format, i.e. whether or not the document is "schema-valid". If errors occur, they are shown in a separate window and the validation process is aborted (see Figure 7). The functionality of the Microsoft XML Parser allows to just display the first encountered error. If the XML document contains more than one error, several correction and validation steps are necessary until no more errors are reported. With the help of the description "Line" and "Line position" it is possible to localise and eliminate the error in the XML document. The commercial software XMLSpy (www.xmlspy.com) facilitates the editing of XML documents. If only corrections are to be made in XML documents the use of free text editors (such as <http://www.notetab.com/>) is sufficient.

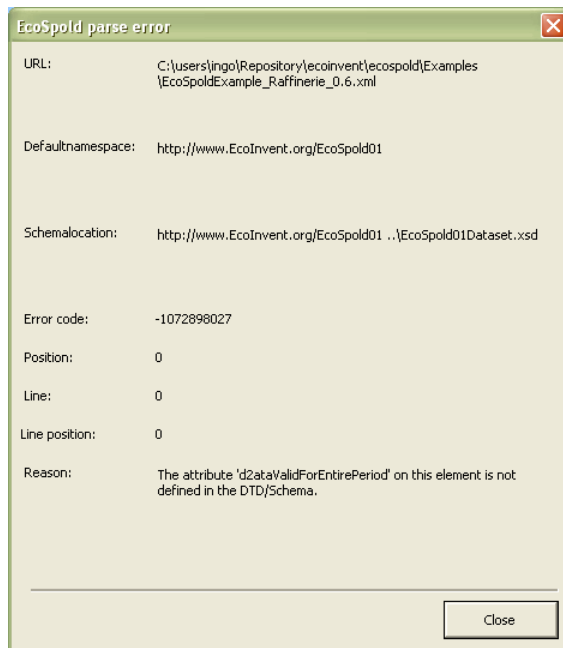


Figure 7: EcoSpold Access: Dialogue of an error of the syntactical validation

If no syntactical error occurs, the EcoSpold document is "schema-valid" and the EcoSpold Access component continues with a semantic validation. If no error occurs a short message is given. If errors occur, a dialogue lists all errors, warnings and hints for each individual dataset of the EcoSpold document (see Figure 8).

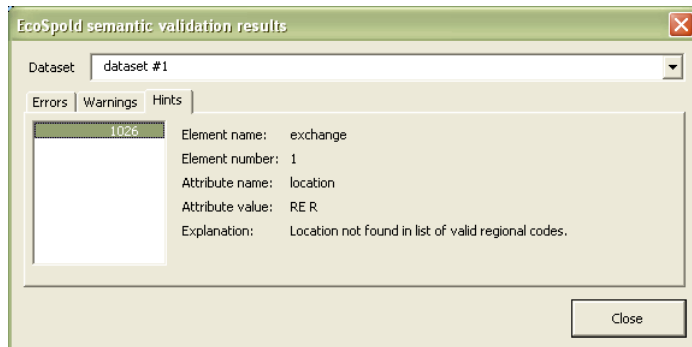


Figure 8: EcoSpold Access: Result dialogue of a semantic validation

The "Dataset" dropdown list in the upper part of the dialogue allows to switch between individual datasets of the respective EcoSpold document. In the lower part, the errors, warnings and hints of the dataset chosen are displayed. If information is available, element name, number, the name of the respective attribute, its value and a comment is reported.

8.6.1 Primary/Foreign key violations

The message of the XML Parser "The keyref "12" does not resolve to a key for the Identity Constraint '{http://www.Ecoinvent.org/EcoSpold01}pkPersonNumber'." means that the value '12' does not comply with a foreign key, which refers to a primary key (pk) 'pkPersonNumber'. The primary key pkPersonNumber defines the set of valid values from the @number attribute of the person elements. Foreign keys refer to this set and enable the XML Parser to determine whether or not a given value is valid. In this case the validation fails because the value '12' is not included in the valid set of person numbers ("does not resolve to a key...").

8.6.2 Schema documentation

The six buttons for the display of the EcoSpold Schema documentations create HTML files in the 'Doc' subdirectory of the EcoSpold path. The print version (View simple...) is identified with „Simple“ at the end of the file name. After the creation of the HTML files they are displayed in the Internet Explorer. Documentations can be created for the three schemas 'Process', 'Elementary Flow' and 'Impact Category'. These documentations are already available in HTML format in the 'Doc' directory.