
Author	Gabor Doka, Doka LCA, Zurich, do@doka.ch
Title image	Excerpt of the EcoSpold01 schema definition from (Ecoinvent 2002)
Thanks	Thanks to Mischa Zschokke, Project Manager Environmental Consulting, carbotech Zürich, for testing the trial version of the converter in September 2022.
Disclaimer	The information contained in this report were carefully elaborated and compiled by the author according to best knowledge and available data. Due to the broad scope of the task and the inherently variable nature of the subject, errors or deviations cannot be excluded. For this reason the information contained in this report is provided without warranty of any kind, express or implied, including but not limited to the fitness for a particular purpose. In no event shall the author, Doka Life Cycle Assessments or the commissioner be liable for any claim, damages or other liability, arising from the use, application or dissemination of this information.
Copyright	<p>Creative Commons Attribution-ShareAlike 4.0 (CC BY-SA 4.0) https://creativecommons.org/licenses/by-sa/4.0/</p> <p>Copyright © 2002-2023 Gabor Doka, do@doka.ch . You are free to:</p> <p>Share — copy and redistribute the material in any medium or format</p> <p>Adapt — remix, transform, and build upon the material</p> <p>Under the following terms:</p> <p>Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.</p> <p>ShareAlike — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.</p> <p>No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.</p> <p>The reports are provided free of charge to anybody. The reports or derivatives may not be sold or offered for sale, or included with software products offered for sale. If you paid money to obtain these reports, somebody has profited illegally from content provided freely. Companies that distribute documents for profit are expressly prohibited from distributing the reports. The reports come with absolutely no warranty. These are free documentations, and you are welcome to redistribute them under the Creative Commons Attribution-ShareAlike conditions.</p>
Declaration	This report and all previous reports by Doka LCA were created without any use of generative artificial intelligences like ChatGPT or similar. Used were however traditional English-German dictionaries and thesauri.

Suggested citation: Doka G. (2023) The Esixell converter for EcoSpold1 process inventories between Excel workbooks and XML files. Doka Life Cycle Assessments, Zurich, Switzerland. Commissioned by Swiss Federal Office for the Environment (FOEN), Berne, Switzerland. April 2023. Available at <http://www.doka.ch/publications.htm>

Significant digits Figures in this report often feature several digits. This is not to imply that all the shown digits are really significant or that the data displayed is very precise. Showing several digits helps to minimise the avoidable accumulation of rounding mistakes along the chain of calculations performed here, and in possible future studies referring to this data.

Percent is not a unit A value like 100% is mathematically identical to 1, and "33%" is just a way to write the value 0.33 (which one could also write in yet another different format as $3.3 \cdot 10^{-1}$). Mere *formatting* does not and should not influence the *magnitude of a value*. There is therefore no need to introduce factors or divisors of 100 in formulas for percentages. "Per cent" literally means "per one hundred" and implies the instruction "divide by 100", therefore the mathematical value of the expression "33%" is $33/100 = 0.33$ (not 33). In contrast, a formula to calculate a gram value from kilograms must include a factor of 1000, because gram is a *physical unit* (and not a different way to "format" a kilogram value). The EcoSpold1 Definition unfortunately has an ambiguous understanding of percentages. In the field for Geometric Standard Deviations², values of 1.5 are considered to be identical to 150%. In the field for market representativeness 1.5 would mean 1.5%, and 100 would mean 100%, i.e. in that latter field, percent is incorrectly treated as a physical unit.

Contents

1	Introduction	6
2	Basic operation	8
2.1	Excel to XML (XL to ML).....	8
2.2	XML to Excel (ML to XL).....	9
2.3	Other macros the user can run.....	10
3	User preference settings.....	11
3.1	Overall preferences	11
3.2	Preferences for XL to ML.....	12
3.3	Preferences for ML to XL.....	12
4	Helper Macros for Editing	14
5	Tips for a properly formatted Excel start file	16
5.1	X-Exchange	16
5.2	X-Process	19
5.3	X-Source.....	22
5.4	X-Person.....	23
5.5	Consecutive ranges for conversion	24
5.6	Colour and formats	25
5.7	Time period fields	28
5.8	Synonyms.....	28
5.9	Review comments	28
5.10	Attribute 'Infrastructure Included'	29
5.11	Automatic defaults for missing cell contents.....	31
6	Conversion ES1-XML to Excel	32
6.1	Content in generated Excel	32
6.2	Warnings and errors from ML to XL	40
6.3	Tips for the start XML file.....	45
7	Conversion Excel to ES1-XML.....	47
8	Troubleshooting Esixell	50
8.1	Where are the converted files?.....	50
8.2	Can't open XML File (ML→XL).....	50
8.3	Can't open Excel start file (XL→ML).....	50
8.4	Esixell toolbar still visible after closing Esixell	50
9	Outlook.....	52
10	Appendix A: Meaning of EcoSpold1 integers	53

10.1 In sheet 'X-Exchange'	53
10.2 In sheet 'X-Process'	53
10.3 In sheet 'X-Source'	54
10.4 In sheet 'X-Person'	54
11 References	55

1 Introduction

The Esixell Converter is a converter for process inventories in EcoSpold1 format. It converts between a source file and a target file format. Conversions from a source Excel file (.xlsx, .xls etc.) to a target EcoSpold1-XML file (.xml text) is possible, or from a source EcoSpold1-XML file to a target Excel file. Since the Excel file contains a sheet named 'X-Source', to avoid confusion the term 'source file' is replaced with 'start file' in this manual.

The goal of Esixell is to partly replace the now deprecated EcoSpoldAccess Add-in for Excel, which converted between EcoSpold1-XML and Excel files (Hedemann & Meinshausen 2003). Esixell only converts *process inventories*. It does not have the full range of functionality of EcoSpoldAccess, as the latter also wrote files for exchange definitions ("NamesElementary") and LCIA methods ("X-ImpactFactor").

Esixell can be used to assist creation of *new* inventories, by compiling inventory data in an Excel file and then convert it to an EcoSpold1-XML file, see Fig. 1.1. Or Esixell can also be used to conveniently *edit* existing inventories by converting XML files to Excel, augment or edit the information, and convert the Excel file back to an EcoSpold1-XML file, see Fig. 1.2.

The Esixell Converter is implemented as an Excel Workbook with macros (Visual Basic VBA). To use Esixell the Microsoft Excel software is therefore necessary and execution of VBA macros need to be allowed in MSOffice/Windows.

For abbreviation in this manual, the XLSX files are abbreviated with XL (for 'Excel'), and the EcoSpold1-XML files are abbreviated with ML (for 'Markup Language'). So the Esixell Converter performs conversions from XL to ML, or from ML to XL.

The name Ésixell stands for leetspeak ES1-XL, which in turn abbreviates EcoSpold1 and Excel.

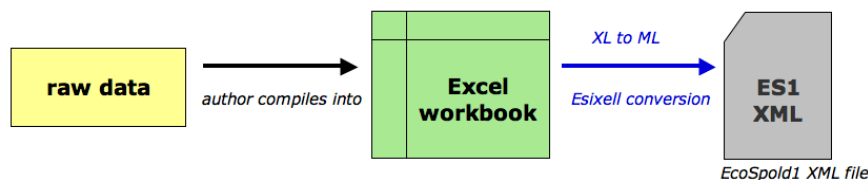


Fig. 1.1 Workflow for generating new process inventories, involving Esixell Converter.

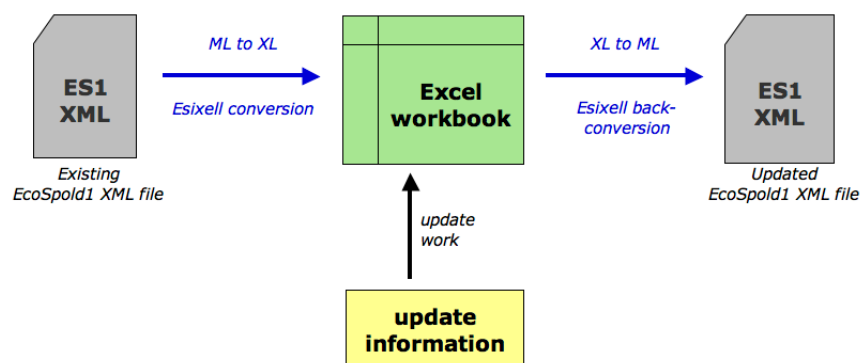


Fig. 1.2 Workflow for updating or extending existing process inventories, involving Esixell Converter.

When converting ML to XL Esixell will try to extract as much information as possible, even from a seriously defective XML file, but highlight detected errors, so that required error correction can be performed by the user in Excel and not in an unwieldy XML text file. For the reverse XL to ML, Esixell will halt conversion if important, mandatory information is missing, as to avoid creating any non-conform XML output files.

Esixell is merely a *converter* between file formats. It is not a *teaching tool* for novices to the EcoSpold1 format. Users must be already proficient in understanding the contents and meanings of the various fields in EcoSpold1.

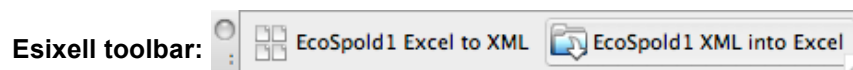
2 Basic operation

Allowing Macro Execution

The Esixell Converter is contained in the Excel Workbook 'EsixellConverter2023.xlsm'. This contains executable code as macros (Visual Basic VBA). To use it, Microsoft Excel software needs to be available¹ and execution of macros needs to be allowed by the user. Some security settings especially on Windows computers are suspicious of documents with macro content. So you need to tell Office/Windows that execution of these workbooks is OK.

Esixell Toolbar/Executing macros

If your MSeExcel version allows, a custom Esixell toolbar is created upon opening the Esixell workbook and displayed on the screen. The toolbar contains buttons for the two main operations of Esixell. Upon closing of the Esixell workbook, this toolbar will be deleted again. In some Excel versions the toolbar will only be removed when Excel is quit entirely.



In Excel versions like Excel 365 automated creation of toolbars is sadly not possible anymore and here macros must be executed by the user with the command ribbon, on tab "Developer"² and choosing the "Macros" icon. This then lists all available macros in your Excel application. At the bottom of the dialog, limit the display to macros contained in the Esixell workbook. The Esixell macros pertinent to the user are listed on top (alphabetically) starting with deliberate prefixes like "Aaaaah_...".

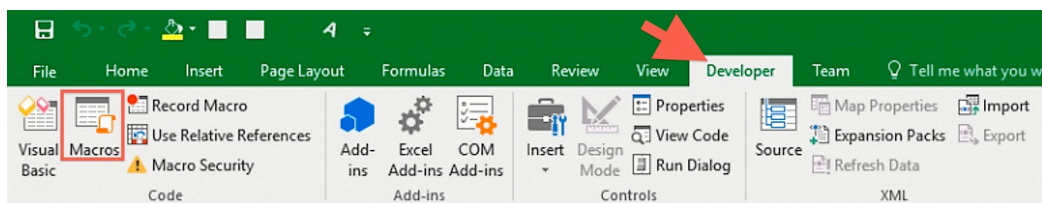



Fig. 2.1 The developer tab in Excel 365. The "Macros" icon is the second from the left.

2.1 Excel to XML (XL to ML)

It is recommended to make first trial runs with a *copy* of the intended Excel workbook file first. Esixell will correct or overwrite certain contents (see combinames in chapter 3.1.2, and defaults in chapter 5.11).

Conversion from a start Excel Workbook to a target EcoSpold1-XML-File (text) involves these steps:

- ¹ Open Source software alternatives like LibreOffice's Calc will not work, because the VBA syntax has not been implemented in Calc. Therefore VBA code will not run on Calc.
- ² The Developer tab might not be visible in your Excel ribbon. You then have to unhide it to display it. To do this in the "File" tab of the ribbon click on the "Options" icon. 2. in the Options dialog box, click "Customize Ribbon". 3. Ensure the Developer option is checked to enable the Developer tab. 4. Click OK to close the Excel Options.

1. **Open the workbook containing the Esixell macros. Allow execution of macros in Office/Windows.**
2. **Open the start workbook containing the properly formatted inventory information (read the note on start files in large folder hierarchies chapter 8.3 on page 50).**
3. **With the start workbook of the previous step in the foreground, press the toolbar button 'EcoSpold1 Excel to XML'  , or run the VBA macro 'Aaaah_XLtoML'³ by choosing it via menu command ('Extras/macros...' or similar) or with the "Macros" icon in the Developer tab of the ribbon.**

This macro is sensitive to the workbook that is in the foreground at the time the macro is called. Make sure the desired ES1-workbook is in the foreground and not some other workbook or the Esixell workbook.

Conversion might take a few seconds or longer, especially for large and/or multi dataset start files. If need be, you can interrupt the macro execution with the 'ESC' key.


Start workbooks might contain single or multiple datasets. These will be converted into one or several generated XML files, depending on the user preferences (cf. chapter 3.2.2 on page 12). The name of the generated XML file is either adopted from the name of the start workbook, or from the names of the individual datasets. If a result file contains only one activity, its name is derived from that activity.

After the conversion is finished a message "Done creating XML" will appear giving details of number of files created and their location.

Esixell *tries* to save its exports into the same folder as the start file. Due to total path length limitations, this might not be possible (start files in deep folder hierarchies cf. chapter 8.3 on page 50). Esixell will then try to *move up* the folder hierarchy automatically, until it can save the file with a short enough file path.

2.2 XML to Excel (ML to XL)

Conversion from an EcoSpold1-XML-File (text) to an Excel Workbook involves these steps:

1. **Open the workbook containing the Esixell macros. Allow execution of macros in Office/Windows.**
2. **Press the toolbar button 'EcoSpold1 XML into Excel'  , or run the VBA macro 'Aaaah_MLtoXL' by choosing it via menu command ('Extras/macros...' or similar) or with the "Macros" icon in the Developer tab of the ribbon.**
3. **In the upcoming dialog, choose an EcoSpold1 XML file as start file for the conversion (read the note on start files in deep folder hierarchies chapter 8.2 on page 50).**

Start XML files might contain single or multiple datasets. These will be converted into one or several generated workbooks, depending on the user preferences (cf. chapter 3.3.1 on page 12). The name of the generated workbook is either adopted from the name of the start XML file, or from the names of the individual datasets. If a result workbook contains only one single activity, its name is derived from that activity.

³ The macro name starts with "Aaaah_" in order to appear at the top of the list of executable macros contained in Esixell.

After the conversion is finished either the created workbook(s) will be shown in Excel, or—in case the user has set preferences to automatic closing of created workbooks—a message "Done creating Excel" will appear giving details of number of files created and their location.

If error(s) are detected in dataset, a summary text will be displayed and activated in sheet 'X-Process' in row 54. Conversely that means if after an import the workbook remains on the first sheet "X-Exchange" that Esixell has not detected any errors in the source XML file.

Esixell *tries* to save its exports into the same folder as the start file. Due to total path length limitations, this might not be possible (start files in deep folder hierarchies cf. chapter 8.2 on page 50). Esixell will then try to *move up* the folder hierarchy automatically, until it can save the file with a short enough path.

2.3 Other macros the user can run

Esixell's principal functionality is the conversion of Excel files into XML and vice versa, as described above. Some supplementary macros are contained in Esixell, which the user can run additionally for certain purposes. This is listed in the table below.

There is no extra toolbar button for these macros. The macros can be run by choosing them from the list of available macros (menu command 'Extras/Macros...' or similar; or ribbon button 'Developer' → Macros... see Fig. 2.1 on page 8). Underneath Excel's macro list the user can select to show only macros from within Esixell.

More details on these helper macros are given in chapter 4 'Helper Macros for Editing' on page 14.

Tab. 2.1 Additional macros in Esixell

Name of macro	Description
Aaah_MakeNewWB	Will create a new workbook with the four required sheets and header names and IDs, but entirely empty inventory content.
Aaae_RemoveSelectedErrorFormattingAndComments	In the currently selected range of cells of the active workbook remove any of Esixell's highlight cell colours (orange or light green) and any cell comments. See chapter 4.1.1 on page 14.
Aaae_RemoveAllErrorFormattingAndComments	In the whole active workbook remove any of Esixell's highlight cell colours (orange or light green) and any cell comments. See chapter 4.1.1 on page 14.
Aaah_AlignAllCommentBoxes	In all sheets of the active workbook all the comment boxes are aligned to the right of the cell they are associated with. See chapter 4.1.2 on page 15.
Aaah_DeleteNumberFormats	Deletes various problematic number formats from the active workbook, i.e. [=0]"";General [=0]"";Standard 0.00E+0;[=0]"";0.00E+0 See section 'Problematic Cell Formatting II' on page 26
Aaah_SetDefaultPreferences	Set the suggested default preferences for Esixell (except for splitting of files). See chapter 3 'User preference settings' on page 11.

3 User preference settings

Esixell contains some preferences which can be set by the user. Preferences are in the Esixell workbook in the sheet 'Preferences'. Preferences are often set by entering 1 or 0 in certain cells for yes/no.

The macro Aaah_SetDefaultPreferences will set the suggested default preferences for Esixell (except for splitting of files).

3.1 Overall preferences

3.1.1 Allow duplicate Exchanges?

Duplicate exchanges within the same dataset can be problematic and are not recommended (see chapter 6.2.1 '[] Unique exchange entries' on page 41). An exchange is defined by its name, unit, its InputGroup or OutputGroup. For biosphere exchanges the category and subcategory is also decisive, and for technosphere flows the location and the boolean for InfrastructureProcess is also decisive. Unfortunately, not every practitioner adheres to this and ES1-XML files can be encountered where identical exchanges are used more than once within the same dataset. Therefore the start file or workbook might contain a particular exchange more than once. In order not to interrupt other people's workflow, Esixell can be set to tolerate duplicate exchanges. Allowing duplicates is not recommended and it is the dataset author's/editor's responsibility to risk having duplicate exchanges in datasets.

The user can choose whether duplicate exchanges shall be allowed (1 = allowed) or be considered an error (0 = not allowed).

During parsing from XL to ML and if duplicates are not allowed, any encountered duplicates will interrupt parsing and result in an error message.

During parsing from ML to XL and if duplicates are not allowed, any encountered duplicate exchanges will be written into an informal repository at the bottom of sheet 'X-Exchange' in blue font. The informal repository is separated by the formal part by an *empty row* and therefore will not be included in a back-conversion into XML. To indicate that it has a blue font.

0 means	duplicate exchanges will be reported as an error
1 means	duplicate exchanges will be included in the formal inventory

Recommended setting is 0.

3.1.2 Keep helper combinames after parsing?

Combinames are helper contents written into in workbooks, helping to find Esixell content and check for duplicate entries. For instance for a literature source a combined string from its first author, publication year and title is formed, e.g. *'Itten R., 2013LCI of global PV supply chain'*. This is the combiname of that publication, which is written into the top row of sheet 'X-Source'. Combinames are written in all four sheets relating either to exchanges, processes, source or persons (usually in the top row, in sheet 'X-Exchange' left of the column 'InputGroup').

These combinames might be helpful for the user's editorial work. The user can choose if the combinames shall be deleted after parsing or kept.

0 means	combinames strings are deleted again after parsing
1 means	combinames strings are kept and not deleted

Recommended setting is 1.

3.2 Preferences for XL to ML

3.2.1 Number of digits in numerical values

Values calculated in Excel can commonly contain 15 digits. Such extent of digits can be justified (e.g. if large amounts are subtracted in downstream processes), but often unnecessary. In the Esixell workbook the sheet 'Preferences' contains a cell entitled 'Round mean values and GSD to number of significant digits' where the user can choose how many significant digits shall be used in writing mean amounts and standard deviation figures. Positive integers and zero are allowed.

0 means	No rounding is applied, and the values are written into XML as they occur in the Excel cell values
..5,6,7..means	Cell values are rounded to this many significant digits

Recommended setting is 8.

3.2.2 Split into several XML files?

The start Excel workbook file might contain a single or multiple activities/datasets. If the workbook contains *multiple activities*, either all datasets can be written into a *single* XML file, or each activity is written into *its own individual* XML file. The user can choose which of these choices is more convenient. In the Esixell workbook the sheet 'Preferences' contains a cell entitled 'Split multiple datasets in start Excel into individual XML files?'. In the white cell either a zero value or 1 can be entered.

0 means	only a single XML file with multiple activities is created, when the start Excel workbook contains multiple activities.
1 means	the contents are split into several XML files each containing only one activity UPR.

Recommended setting depends on number of converted datasets and purpose.

Please note that this setting here is *independent* of the setting for the reverse conversion ML to XL (cf. chapter 3.3.1 below).

If a created XML file contains only a single activity, the file name will be derived from the name of the activity. If an XML files contains multiple activities, the file name will be derived from the name of the start workbook.

3.3 Preferences for ML to XL

3.3.1 Split into several workbooks?

The start XML file might contain single or multiple activities/datasets. If the file contains *multiple activities*, either all datasets can be written into a *single* Excel workbook, or each activity is written into *its own individual* workbook. The user can choose which of these choices is more convenient.⁴ In

⁴ A single workbook with several activities is more convenient for comparison of exchanges and highlighting differences in different activity inventories. But a single workbook also requires consistent ID numbers for referenced literature sources and persons, if more than one is used. With multiple created workbooks, the major drawback is that the list of exchanges in 'X-Exchange' will only contain a sequence of the exchanges used by that activity. Thus different activities will naturally often have different lists of exchanges.

the Esixell workbook the sheet 'Preferences' contains a cell entitled 'Split multiple datasets in XML into individual Excel workbooks'. In the white cell either a zero value or 1 can be entered.

0 means	only a single new workbook with multiple activities is created, when the start XML file contains multiple activities.
1 means	the contents are split into several new workbooks each containing only one activity UPR.

Recommended setting depends on number of converted datasets and purpose.

Please note that this setting here is *independent* of the setting for the reverse conversion XL to ML (cf. chapter 3.2.2 above).

If a created workbook contains only a single activity, the workbook name will be derived from the name of the parsed activity. If a workbook contains multiple activities, the file name will be derived from the name of the start XML file.

3.3.2 Close created workbooks?

After a workbook has been finished and saved, shall it remain open in Excel, or shall it be closed?

If a start XML file contains multiple datasets and each dataset receives its own workbook, the possibly large number of open workbooks can slow down Excel, or at worst crash it. With this preference set to 1 (true) the created workbooks are saved and closed after the parse.

0 means	all created workbooks are saved and remain open in Excel after parsing.
1 means	all created workbooks are saved and closed after parsing.

Recommended setting is 0, unless a great many workbooks are created.

3.3.3 Sort parsed exchanges list?

Esixell reads out the exchanges in the sequence they appear in the start XML file. This can result in unwieldy layouts, especially when more than one dataset is included in a workbook. After the optional sorting, the exchanges are organised alphabetically, with all biosphere exchanges on top, followed by all technosphere exchanges, and reference products at the bottom.

0 means	exchanges in same sequence as in start XML file
1 means	exchanges sorted by kind and alphabetically

Recommended setting is 1.

4 Helper Macros for Editing

During the user's editing and error correction of Excel files some functionalities can be helpful. Esixell contains several additional macros which can help the editing process of Excel files. The macros don't have a separate toolbar symbol, but can be called from the Macro Menu (Extras/Macro/Macro... or similar) with the "Macros" icon in the Developer tab of the ribbon.

Naturally, these macros might *help* with your editing work, but they cannot *resolve* found errors by themselves.

4.1.1 Macros to remove error formatting and comments

Esixell can colour cells to signal to the user errors in content (in orange) or automatically corrected content (in light green). Also often cell comments are added to errors. This shall guide the user to perform required content corrections. After the user has corrected any errors, the cell colours and cell comments are not needed anymore. Although Esixell ignores any cell formats or cell comments during conversion of XL back into ML, it is superfluous—and at worst misleading—to keep those highlights and comments.

The macro 'RemoveSelectedErrorFormattingAndComments' looks in the currently selected range of cells for **cells coloured as orange or light green, and only in those cells removes fill colours and any cell comments attached to them**. This shall be done after all required content changes and checks have been performed by the user. The macro is naturally not able to correct any content errors itself. It only removes the cell colours and cell comments.

The macro 'RemoveAllErrorFormattingAndComments' cleans the *whole active workbook* from error and change formats, but only does it in the sheets "X-Exchange", "X-Process", "X-Source", "X-Person" and not any other sheets.

Parsed file with highlighted errors:

CASNUMBER		
StartDate	1900-01-01	year 1900?
EndDate	1900-01-01	
DataValidForEntirePeriod	1	Original DataValidForEntirePeriod attribute was empty (""). Check if 1=true is correct for this activity.
OtherPeriodText	Undefined	
Text	European Electricity mix and material processing data for supplier outside Europe.	
Text	Production by conventional pulverisation of styrene butadiene copolymer that is produced by emulsion polymerisation.	

**After manually corrected and checked cell content:**

CASNUMBER		
StartDate	1980-01-01	year 1900?
EndDate	2000-12-31	
DataValidForEntirePeriod	1	Original DataValidForEntirePeriod attribute was empty (""). Check if 1=true is correct for this activity.
OtherPeriodText	Undefined	
Text	European Electricity mix and material processing data for supplier outside Europe.	
Text	Production by conventional pulverisation of styrene butadiene copolymer that is produced by emulsion polymerisation.	

**After running 'RemoveSelectedErrorFormattingAndComments' over the pertinent cell range**

CASNUMBER		
StartDate	1980-01-01	
EndDate	2000-12-31	
DataValidForEntirePeriod	1	
OtherPeriodText	Undefined	
Text	European Electricity mix and material processing data for supplier outside Europe.	
Text	Production by conventional pulverisation of styrene butadiene copolymer that is produced by emulsion polymerisation.	

4.1.2 Macro 'AlignAllCommentBoxes'

For ML to XL conversion, Esixell can write comments into cell comment boxes (see for example images above). When Excel tables are edited manually these comment boxes might be shifted into odd positions. This macro aligns all the comment boxes in all sheets to the right of the cell they are associated with.

4.1.3 Macro 'DeleteNumberFormats'

As explained in section 'Problematic Cell Formatting II' on page 26 Excel workbooks can contain problematic user-defined number formats, i.e. formats that display zero values as hyphens ("-") or as nothing. Since Esixell works on the *values* a cell actually contains, not on what is *displayed*, this can be confusing and lead to problems. Even in previously unproblematic workbooks user-defined number formats can be reintroduced by copy-pasting from other sources. Pasting from other workbooks should ideally only be as values without cell formatting.

5 Tips for a properly formatted Excel start file

Templates: When creating an inventory in an Excel table, it is best to work from a proper template. A template can easily be obtained by running the macro 'MakeNewWB' from Esixell.⁵

In the text below some features are pointed out, especially those required for conversion from XL to ML with the Esixell Converter.

Worksheets: The workbook needs to contain *four worksheets* with these following exact names

- X-Exchange
- X-Process
- X-Source
- X-Person

Other sheets can be present, but the four sheets listed above contain the information used for conversion. Please note that all sheet names contain words in singular form, e.g. 'X-Person' not 'X-Persons'.

X-Exchange contains the biosphere and technosphere exchanges and uncertainty information.

X-Process contains the meta information, time period, text comments for the inventoried activity, authorship, indicates literature references etc.

X-Source contains the details to literature sources indicated in the metadata.⁶

X-Person contains the information to any people indicated in the metadata, i.e. their name, contact details and affiliation.

For each inventory listed in X-Exchange, a column in X-Process must exist with contains its metadata. In turn, the metadata might reference sources or persons, which need to be present in columns of X-Source and X-Person.

5.1 X-Exchange

This sheet contains the process inventory as such (biosphere and technosphere exchanges). It has usually a completely empty top row and empty leftmost column (row 1 in Excel and column A). While this is not strictly required for the Esixell Converter, it was required in earlier ES1 Excel templates and to minimise compatibility issues with other sources, this layout choice should be maintained.

⁵ Different templates are also available from various online sources, but they sometimes feature problematic cell formats, see chapter 5.6 'Colour and formats' on page 22.

⁶ To reference a literature source (even one's own) is recommended for transparency. But the EcoSpold01 schema does not require to reference a source in all cases. If in X-Process the field 'dataPublishedIn' (ID 756) contains the values 0="Data as such not published" or 1="The data of some unit processes or subsystems are published", the field referenceToPublishedSource (ID 757) in X-Process *can remain undefined* and in that case the sheet 'X-Source' can remain without any source entries. A source is however required and mandatory, if 'dataPublishedIn' in X-Process is 2 = "Data has been published entirely in 'referenceToPublishedSource' ". Actually, the EcoSpold01 schema remarks that "within theecoinvent quality network all datasets are published in the series of ecoinvent reports" and since EcoSpold01 has no other value for the QualityNetwork (field ID 304) than the value 1 for 'ecoinvent', it might even concluded that in EcoSpold01 published *reports are mandatory*, and therefore no other value than 'dataPublishedIn' = 2 is correct.

Also column B, headed 'ID', is largely empty. It has been used in the past to write titles for groups of exchanges, like "emissions to air". This is however informal. Esixell will use this column below row 6 to write combinames of exchanges as a quick way to find particular exchanges (cf. chapter 3.1.2 on page 11). Thus any previous content will be overwritten.

5.1.1 Anchor range

For better fault tolerance, the Esixell Converter does not rely on the above choices of empty rows/columns—which are still recommended—but it will try to find the top left cell of each of the four inventory sheets by looking for a particular sequence of cells that contain the ID numbers of required fields. This so called 'anchor range' is indicated in blue in the following figures.

For the Exchange table the top left contains the following cells (with Excel row numbers on the left, and Excel column characters on top).

	A	B	C	D	E
1					
2			ID	3503	3504
3		Explanation	401	InputGroup	OutputGroup
4			662		
5			493		
6			403		

This arrangement will be sought out by the Esixell Converter to establish the top left edge of the Exchange table. For this, the anchor range need to be present (ID, 3503, 3504, 401, 662, 493, 403. I.e. the cells shown in blue) and in this precise arrangement. Only numbers and the "ID" cell are heeded and other cells with text are ignored for this: In some templates the phrases "Input-Group" and "Output-Group" (with hyphen) are used, although these are not the proper EcoSpold1 names of those elements. In the shown figures the top left anchor cell of the table content is indicated in green (here cell B2).

Tab. 5.1 Concept of finding the top left cell for inventory content in the worksheets



Tab. 5.2 Example Excel worksheet for 'X-Exchange'

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
2		ID	3503	3504	3702	3703	3506	3507	3508	3706	3707	3708	3709	3795	3796	3797	3792
3		401	InputGroup	OutputGroup	Name	Location	Category	SubCategory	InfrastructureProcess	Unit	disposal, average municipal waste, 22.9% water, to incineration	UncertaintyType	StandardDeviation95%	MinValue	MaxValue	MostLikelyValue	GeneralComment
4		662			Location						CH						
5		493			InfrastructureProcess						0						
6		403			Unit						kg						
7				4	Carbon monoxide, biogenic		air	high population density		kg	1.34207E-05	1	507.0%				(1,4,1,1,3,1)
8				4	Carbon monoxide, fossil		air	high population density		kg	6.5793E-06	1	507.0%				(1,4,1,1,3,1)
9				4	Carbon dioxide, biogenic		air	high population density		kg	0.517545416	1	143.6%				calculated

In the Exchange table there is one row for each exchange (row 7 and below in the above example). Individual processes are in separate columns. What is contained in each column is written in row 3 as a horizontal header text. At the beginning of each new process column, the process is identified by giving its name, location code, infrastructure flag and physical unit (indicated in the table above with a yellow cell fill colour). The titles for those header entries is given in the cells just below the ID number 3702 (column F in the above example). The top left area of X-Exchange therefore by tradition contains a combination of horizontal header text (for exchanges) and vertical header text (for processes).

In the row 2, directly above the horizontal header text, the ID number of that field is given. The presence of the correct ID number at that location is vital for the accurate functioning of the Esixell Converter. Without those ID numbers for the field the inventory cannot be converted.

5.1.2 Sequence of cells

The exchange identification in 'X-Exchange' has eight columns which must appear in a particular sequence.

3503	3504	3702	3703	3506	3507	3508	3706
InputGroup	OutputGroup	Name	Location	Category	SubCategory	InfrastructureProcess	Unit

Esixell will use the ID numbers to check whether the layout is correct. The names in the row underneath are *not* checked by Esixell.

The column for 'Name' will be used to determine the consecutive rows of listed exchanges (cf. chapter 5.5 'Consecutive ranges for conversion' on page 24). Thus the following vertical header texts are required in that column as well, without blank cells:

3702
Name
Location
InfrastructureProcess
Unit

For each dataset/activity in 'X-Exchange' further columns are required. Also here the ID number is used to check whether the layout is correct, and name text strings are ignored.

3707	3708	3709	3795	3796	3792
[meanValue]	UncertaintyType	StandardDeviation95%	MinValue	MaxValue	GeneralComment

Of these columns, the columns for MinValue and MaxValue can be omitted entirely, but only when neither triangular nor uniform distributions are used in the inventory data of an activity. However, if they are used, they ideally appear in the sequence displayed above (first Min, then Max, between StandardDeviation95% and GeneralComment). In an Excel file with multiple datasets, the columns for MinValue and MaxValue must only be present in those activities which require them, and can be omitted in the others. Esixell will determine the required columns for multiple datasets in a flexible manner.⁷

5.2 X-Process

This sheet contains what is sometimes referred to as the meta information: overarching information about the inventoried process. Comments on technology, geography, time period, data sources used, authorship etc.

5.2.1 Anchor range

The values and arrangement of the cells in blue in the figure below are used to find the top left cell in this worksheet, and must be present in this fashion. In the shown figures the top left cell anchor is indicated in green.

	A	B	C
1	Type	ID	Field name
2	ReferenceFunction	401	Name
3	Geography	662	Location
4	ReferenceFunction	493	InfrastructureProcess
5	ReferenceFunction	403	Unit

The row 1 (headed 'Type') is commonly empty after of column C. Esixell will use this row to write out combinames of processes as a quick way to find a particular dataset's metadata (cf. chapter 3.1.2 on page 11). A dataset's combiname is derived from its Name+Location+Infrastructure+Unit. Thus any previous content in that row will be overwritten.

5.2.2 Sequence of cells

The sheet 'X-process' contains several fields with meta information for each activity dataset. For each activity a column must be present.

⁷ A column for the Maximum Likelihood Value (MLV) of a triangular distribution is *not* required. The attribute is defined in the EcoSpold01 schema definition (mostLikelyValue, #3797), but together with a comment *not* to use it. In a triangular distribution, the presence of min, mean, and max values implicitly defines the MLV, so any MLV only duplicates information. Esixell will ignore any MLV columns in XL → ML, but will parse MLV, if it is encountered in ML → XL.

Esixell will use the ID numbers of all rows in the header to check whether the layout of 'X-Process' is as expected (in the figure above the ID numbers are in column B). However, the name strings in the column to the right are not checked by Esixell. The proper sequence of fields and attributes are as shown below. All rows and especially the ID numbers are necessary to be present *in this exact order*, even if they are not mandatory attributes of EcoSpold1 or left empty for a dataset.

Nominally mandatory attributes of the EcoSpold1 format are printed in bold below. The validation attributes 'proofReadingDetails' and 'proofReadingValidator' would nominally be mandatory too, but were decided to be optional content in Esixell, see explanation in chapter 5.9 'Review comments' on page 28.

Row	Type	ID	Field name	Description
2	ReferenceFunction	401	Name	Name of activity
3	Geography	662	Location	Location (ISO code)
4	ReferenceFunction	493	InfrastructureProcess	0/1 for false/true
5	ReferenceFunction	403	Unit	physical unit (f.u.)
6	DataSetInformation	201	Type	1 = unit process inventory
7		202	Version	database version ¹
8		203	energyValues	always 0 (meaningless)
9		205	LanguageCode	usually = "en"
10		206	LocalLanguageCode	usually = "de"
11	DataEntryBy	302	Person	number indicating person
12		304	QualityNetwork	always 1 (=ecoinvent)
13	ReferenceFunction	400	DataSetRelatesToProduct	always 1 (= true)
14		402	IncludedProcesses	Text description for the inventory
15		404	Amount	Always 1 (f.u.) → cf. ID #403
16		490	LocalName	Name of activity in LocalLanguageCode (usually German)
17		491	Synonyms	optional synonyms for the activity or output material or service
18		492	GeneralComment	Text description for the inventory
19		494	InfrastructureIncluded	usually 0 (=false). See chapter 5.10 on page 29.
20		495	Category	Category of this activity, e.g. "waste management"
21		496	SubCategory	Subcategory of this activity e.g. "municipal incineration"
22		497	LocalCategory	Category in LocalLanguageCode
23		498	LocalSubCategory	Subcategory in LocalLanguageCode
24		499	Formula	chemical formula, e.g. "C2H4" for Polyethylene
25		501	StatisticalClassification	²
26		502	CASNumber	with hyphens, eg. "57-12-5"
27	TimePeriod	601	StartDate	as "yyyy-mm-dd"
28		602	EndDate	as "yyyy-mm-dd"
29		603	DataValidForEntirePeriod	usually 1 (= true)
30		611	OtherPeriodText	Text comment
31	Geography	663	Text	Text comment
32	Technology	692	Text	Text comment
33	Representativeness	722	Percent	0-100, market supply
34		724	ProductionVolume	Annual consumption in area (sic!) ³
35		725	SamplingProcedure	Text comment
36		726	Extrapolations	Text comment
37		727	UncertaintyAdjustments	Text comment
38	DataGeneratorAndPublication	751	Person	number indicating person
39		756	DataPublishedIn	code number ⁴
40		757	ReferenceToPublishedSource	number indicating source
41		758	Copyright	usually 1 (=true)
42		759	AccessRestrictedTo	usually 0 (=open)
43		760	CompanyCode	only for restricted LCIs
44		761	CountryCode	only for restricted LCIs
45		762	PageNumbers	refers to source (ID #757)
46	Validation	5615	proofReadingDetails	Review text comment
47		5616	proofReadingValidator	number indicating person
48		5619	otherDetails	Further feedback

1 Format "#.##". For datasets actually having been part of the ecoinvent database, values of "1.00" up to "2.20" are possible. For all other datasets no such number is appropriate and "0.00" should be entered.

2 According to the EcoSpold01 schema definition, this should be the EU classification NACE code (Nomenclature of Economic Activities) but is not really used in all of ecoinvent. Some third parties use SITC (Standard International Trade Classification).

- 3 According to the EcoSpold1 format definition the field productionVolume "indicates the market area consumption volume (NOT necessarily identical with the production volume) in the geographical area indicated of the product/service at issue. The market volume should be given in absolute terms per year and in common units. It is related to the time period specified elsewhere." The bracketed phrase is in the original text (Ecoinvent 2002).
- 4 Usually 2 is used. Possible values: 0= Data as such not published elsewhere, 1= The data of some unit processes or subsystems are published, 2= Data has been published entirely in (reference to number in ReferenceToPublishedSource)

5.3 X-Source

5.3.1 Anchor range

The values and arrangement of the cells in blue in the figure below are used to find the top left cell in this worksheet, and must be present in this fashion. In the shown figures the top left cell anchor is indicated in green.

	A	B	C
1	Type	ID	Field name
2	Sources	801	Number
3	Sources	802	SourceType

The row 1 (headed 'Type') is commonly empty after column C. Esixell will use this row to write out combinames of sources as a quick way to find a particular source (cf. chapter 3.1.2 on page 11). A sources's combiname is derived from its First Autor+Year+Title. Thus any previous content in that row will be overwritten.

5.3.2 Sequence of cells

The full source info contains more cells than just the anchor range shown above. Esixell will use the ID numbers of all rows in the header to check whether the layout of 'X-Source' is as expected (in the figure above the ID numbers are in column B). However, the name strings in the column to the right are not checked by Esixell. The proper sequence of fields and attributes are as shown below. All rows and especially the ID numbers are expected and necessary to be present *in this exact order*, even if they are not mandatory attributes of EcoSpold1 or left empty for a source.

Below the expected sequence of cells is shown together with a short description. For a complete description and usage see (Ecoinvent 2002). Names in bold print indicate mandatory content, while the other cells contain optional information (but must be present even if empty).

Row	ID	Name	Contents	Example
2	801	Number	Identifying integer number for this source	1
3	802	SourceType	Number for type of source *	2
4	803	Text	Comment/description of the source, summary	CD-ROM
5	1002	FirstAuthor	Name of first author	Nemecek, T.
6	1003	AdditionalAuthors	Name(s) of further author(s)	Kägi, T., Blaser, S.
7	1004	Year	Year of publication	2007
8	1005	Title	Title of publication	Life Cycle Inventories of Agricultural Production Systems
9	1006	PageNumbers	Point to relevant content especially in longer publications	p.52
10	1007	NameOfEditors	Names of editor(s)	
11	1008	TitleOfAnthology	Title of incorporating anthology	Ecoinvent report version 2.0
12	1009	PlaceOfPublication	Company providing information. Also web-address (URL) for reports	ART
13	1010	Publisher	Name of publishing company	Swiss Centre for LCI
14	1011	Journal	Title of publishing Journal	
15	1012	VolumeNo	Volume number of Journal	
16	1013	IssueNo	Issue number of Journal	

* For source type the used numbers are: 0=Undefined (default); 1=Published Article in Journal; 2=Chapters in anthology; 3=Separate publication; 4=Measurement on site; 5=Oral communication; 6=Personal written communication; 7=Questionnaires.

5.4 X-Person

5.4.1 Anchor range

The values and arrangement of the cells in blue in the figure below are used to find the top left cell in this worksheet, and must be present in this fashion. In the shown figures the top left cell anchor is indicated in green.

	A	B	C
1	Type	ID	Field name
2	Persons	5800	Number
3	Persons	5802	Name
4	Persons	5803	Address
5	Persons	5804	Telephone

The row 1 (headed 'Type') is commonly empty after column C. Esixell will use this row to write out combinames of persons as a quick way to find a particular person (cf. chapter 3.1.2 on page 11). A person's combiname is derived from its Name+Company Code. Thus any previous content in that row will be overwritten.

5.4.2 Sequence of cells

The full person info contains more cells than just the anchor range shown above. Esixell will use the ID numbers of all rows in the header to check whether the layout of 'X-Person' is as expected (in the figure above the ID numbers are in column B). However, the name strings in the column to the right

are not checked by Esixell. The proper sequence of fields and attributes are as shown below. All rows and especially the ID numbers are expected and necessary to be present *in this exact order*, even if they are not mandatory attributes of EcoSpold1 or left empty for a person.

The table below shows the expected sequence of cells together with a short description. For a complete description and usage see (Ecoinvent 2002). Names in bold print indicate mandatory content, while the other cells contain optional information (but must be present even when left empty).

Persons are indicated for several purposes: as dataset authors, as data entry persons, or as dataset reviewers. A person is listed together with its employing organisation. It is therefore possible to have the same person listed *more than once* when content in older datasets reflects a *previous* employment status. It is not advised to change this information to achieve unique person entries.⁸

Row	ID	Field name	Contents	Example
2	5800	Number	Identifying integer number for this person	3
3	5802	Name	Full name of person	Gabor Doka
4	5803	Address	Complete postal address	Stationsstrasse 32, 8003 Zurich, Switzerland
5	5804	Telephone	Telephone number	044 463 16 08
6	5805	Telefax	Fax number	
7	5806	Email	email address	gabor@doka.ch
8	5807	CompanyCode	7 letter company code	DOKA
9	5808	CountryCode	2 letter ISO-country code	CH

5.5 Consecutive ranges for conversion

The Excel sheets can contain *multiple* entries, i.e. multiple process inventories (X-Exchange) with their metadata (X-Process) and referenced sources (X-Source) and persons (X-Person).

In all 4 sheets the individual entries must be in **consecutive columns**. No empty columns or rows between entries are allowed. This also means that anything right of an *entirely* empty column will be ignored in a table (and can be used for internal comments or additional calculations, if desired). To create such a break, a totally empty column must be inserted or maintained. In 'X-Exchange' especially the cells in row 2 with the field's ID number must remain empty to create such a break. The uninterrupted, consecutive contents of row 2 in 'X-Exchange' will be used to guide the converter.

Also entirely empty rows will form a delimiter for parsing. Anything below an *entirely* empty row, won't be used in a conversion. In 'X-Exchange' the rows containing the exchange name will be used to determine consecutive rows (column F in the figure above with ID 3702), since other fields could contain non-consecutive entries. In the other the sheets, the second column with the ID numbers is used by the converter to detect breaks in consecutive rows.

⁸ For instance if a person of a private consulting firm had reviewed a dataset in the past, but now this person works in the national Environmental Protection Agency, it is misleading to update the change of this person's of employment in the *old* dataset's reviewer info. This would make it seem that this dataset has "being reviewed by the nation's EPA", when it was not. Also in the reverse situation an update is not appropriate and would distort information.

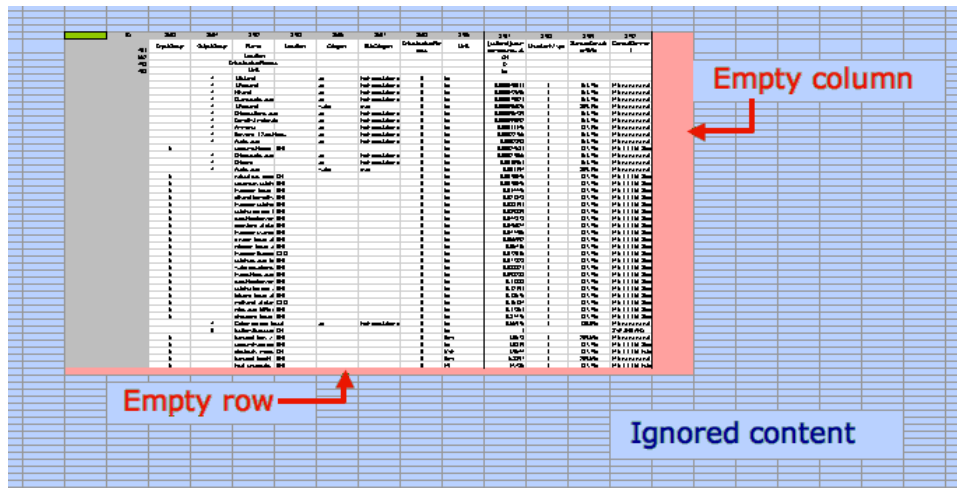


Fig. 5.1 Any content of the Excel tables separated by an empty row and empty column (here highlighted red) is ignored by Esixell when creating XML files. Also anything above or left of the identified top left cell of the table (indicated in green) is ignored.

Esixell's parsing starts by identifying the respective top left header cells in each four Excel sheets (shown as green cell in figure above. See also chapters above). So empty columns *left* of those header cells will not affect parsing, nor will empty rows *above* them. These are also informal parts of the worksheets, disregarded by the XL→ML conversion. Though for maximum compatibility with other sources, it is advised not to introduce more than the usual additional rows above or columns left of these header cells.

As explained above, each inventory listed in X-Exchange required a corresponding column in X-Process and the metadata might reference sources or persons, which need to be present in columns of X-Source and X-Person. It is possible to have *surplus* X-Process columns, that are not actually required in X-Exchange. It is possible to have surplus X-Source and X-Person columns, that are not called upon by required X-Process columns. Esixell will only parse the actually required information in this succession into an ML file. **The content of the surplus columns is not checked for validity and conformity.** It is therefore possible that fields in surplus columns contain errors and incompatibilities. A flawless conversion is not a guarantee that *every* field has been checked, only the actually required information in the defined activities/processes has been checked for conformity and parsed.

5.6 Colour and formats

Generally, the Esixell Converter ignores any cell fill colours or text colours, text or number formatting, or Excel formulas. It converts the *values* contained in the cells. Also Excel's cell comment boxes are ignored. Therefore, cells can have any colour or formatting, but need to be in the right *arrangement and sequence*.⁹ Excel's merging of multiple cells into a single large cell shall not be used.

⁹ Esixell can write some accompanying information in an informal part on the bottom of sheets 'X-Process' and 'X-Exchange' in blue font colour. The blue font is merely a visual highlight for the benefit of the user and exerts no control over content, i.e. giving a part of the *formal* inventory a blue font colour will not lead to Esixell ignoring that part. The delineation against the formal part of the inventory is described in chapter 5.5 on page 21.

Problematic Cell Formatting I: Excel's automatic date conversion

In the standard Excel cell format, Microsoft Excel will try to recognize *dates and times* and convert them automatically into Excel's own date number. E.g. a typed in entry like '12-01' will be *changed* by Excel into a *number value* for an Excel date 42'746 and *displayed* in Excel as the date '12. Jan 2022'¹⁰. After Excel's automatic date conversion, the original string '12-01' actually entered by the user is not retrievable anymore!

This automatic feature of Excel can disrupt manual data entry or editing, e.g. when CAS numbers like 1985-27-9 or page number ranges like 5-12 are converted to date numbers. Since this automatic feature of Excel cannot be switched off in the Standard format, the solution is to format pertinent cells as *text*. If a cell in Excel is formatted as text, Excel will not attempt a conversion to a date number and the user's intended content is conserved.

Following rows are formatted as text to prevent conversion to a date number.

X-Process	Statistical Classification, CAS Numbers, Start and End Dates ¹ , Version number
X-Source	Page numbers, Issue number
X-Person	Telephone, Telefax

¹ Start and End Dates of the time period need to be entered as strings with a distinct EcoSpold1 format, i.e. 'yyyy-mm-dd' which should not be converted by Excel into date numbers or differently formatted strings based on the user's operation system settings.

Problematic Cell Formatting II: used-defined number formats

In templates for EcoSpold1 workbooks sometimes problematic number formatting can be encountered. One such number format is this:

[=0]"-";Standard

This is a user-defined number format and can be contained in a workbook as an available number format of a cell range. The example format above checks if a cell value is zero [=0] and if so displays a hyphen instead: "-". For all other values the standard number format is used. In cells with this format, zero values are *displayed* as hyphens. It is important to understand that this is a mere effect on appearance. Not the cell *value* is changed—the cell value remains zero—but instead of a zero a hyphen is *shown*.

An even more treacherous number format is:

[=0]"";Standard

Here if a cell value is zero, *nothing* is displayed on screen.

While these formatting choices might be helpful in creating uncluttered, nice-looking tables for publications, for the primary task of defining EcoSpold1 content, it can be very misleading, if a value

¹⁰ The used date display will depend on country settings for the operation system and Excel's settings. Excel's date number is usually number of days since 1. Jan 1904.

of zero is displayed as nothing.¹¹ Since the Esixell Converter retrieves actual cell values, and does not look at their mere display, this can lead to misunderstandings and unhelpful content in XML files.¹²

Remedy for problematic user-defined cell formatting

Excel workbooks created by Esixell will not contain such problematic user-defined formats. But Excel templates from other sources might contain them. It is advised to check for such formats (with the command `Format.../Cells.../Number`) and *remove them* from the templates altogether. This can be achieved either manually¹³ or by running following specific macro in Esixell. Esixell contains an optional macro that will delete unwanted number formats shown above and similar ones from the active workbook.

To use it:

- a) Open the Esixell workbook and allow macro execution
- b) Open the target workbook to be cleaned
- c) With the target workbook in the foreground, run the macro 'DeleteNumberFormats' from Esixell
- d) Save the target workbook.

For nice-looking tables to be put in reports, it is advised to create automatic copies of cell ranges, and change unwanted content with formulas instead of cell formats. E.g. a formula like `'=IF(D7=0;"-";D7)'` will convert any zero values in cell D7 to a hyphen value (string). And else simply 'D7' will reproduce the cell value. In rows below the actual EcoSpold content—and separated from the actual content by a completely empty row—such a range can reproduce the information entered above in the desired fashion, and would automatically adapt, if that information were edited. Inserted rows in the inventory would however also have to be inserted into this separated display table.

Hidden ranges

The Esixell Converter will unhide and ungroup any rows or columns in all four pertinent worksheets. This guarantees that the required cell contents, especially in header cells, can be found by the Esixell macros.

¹¹ For instance in the column OutputGroup in sheet 'X-Process' a value of *zero* defines an exchange as the reference product, while an *empty* cell is required for all input exchanges.

¹² For instance if the synonym field is *apparently* empty, but *actually* contains a zero value, previous EcoSpold converters would create an unhelpful and in fact unintended tag `<synonym>0</synonym>`. In Esixell values are for the most part checked for sensible entries, and synonyms with a text length of 1 character are ignored.

¹³ For manual deletion open the template file (or EcoSpold Excel file), click on a cell, choose the "Cell Format..." command (in "Formats/Cell" or similar), choose the 'numbers' tag, selecting 'user-defined formats', and scroll through the list of formats, click on a format similar to the above ones and click 'delete'. The formats are irretrievably deleted, but this affects only the display of cells, not their actual content.

5.7 Time period fields

For the time period of an activity the EcoSpold1 format defines two separate tags that indicate the start date and the end date, which are aptly named `<startDate>` and `<endDate>`. Though usually datasets have time periods of one or several years, they can be shorter than a year—e.g. for seasonal inventories—or could theoretically even pertain to one single day. According to the EcoSpold1 format definition, the time period fields can either contain a year value (yyyy), a year-month value (yyyy-mm), or a full date (yyyy-mm-dd). The hyphen ("-") in these text strings is a common minus sign (not an endash, emdash, non-breaking hyphen or other).

In some encountered XML files from other sources, this change in content length is accompanied by a change of the *name of the used tags*. I.e. year values are given with the tag named `<startYear>` and `<endYear>` and year-month values are given with the tag named `<startYearMonth>` and `<endYearMonth>`, although those tag names are non-canonical, i.e. not really defined in the EcoSpold1 format.

When converting from XL to ML, Esixell will write out the canonical tags `<startDate>` and `<endDate>`, and will always use a *full date* as (yyyy-mm-dd). If the author only specifies a year in the start date cell, it will be expanded to (yyyy-01-01). If only a year and month is given, it will be expanded to (yyyy-mm-01). Similarly for the end date cell, a year value will be expanded to (yyyy-12-31) and a year and month value to (yyyy-mm-31).¹⁴

5.8 Synonyms

In sheet 'X-Process' in field ID #491 optional synonyms for the activity and/or output material or service can be given. Several entries are possible. Also popular abbreviations or alternative spellings can be given, as well as entries in several languages. The separating delimiter is `//`, two normal slash characters¹⁵. From this Esixell will generate separate synonym entries in the generated XML file, by looking for this delimiter. A possible cell value defining synonyms could be for example "rock wool//mineral wool//Mineralwolle//Steinwolle". Please note there are no delimiters at the start or at the end. Also no carriage returns or line feeds should be used in this cell content.

5.9 Review comments

The EcoSpold1 format definition contains a tag `<validation>` which carries a review text and identifies the person that reviewed the dataset.¹⁶ The three validation attributes are displayed in the lower rows of the sheet 'X-Process'.¹⁷

¹⁴ Esixell will actually check the given month (mm) and apply the appropriate number for the last day of that month. In most cases the month will be December and the last day the 31th. But if the month were to be February, the last day actually also depends on the given year (29th in leap years). Esixell will determine the correct last day of the indicated month *and* year.

¹⁵ The `//` character, commonly present in keyboards and colloquially known as "(forward) slash", is formally named "solidus" (codepoint 2F in ASCII, ANSI, or Unicode).

¹⁶ The tag contains three attributes 1) `proofReadingDetails`: Comment text of the reviewer, 2) `proofReadingValidator`: a number identifying the reviewer, 3) `otherDetails`: further text added later, e.g. feedback from users. The first two items are formally mandatory (Ecoinvent 2002)

¹⁷ Usually rows 46 – 48, in *black* font colour. The *blue* font colour indicates an *informal* repository, which is disregarded during back-conversion to ML

The dataset author is not—or should not be—the same person as the dataset reviewer. According to the EcoSpold1 format definition, a review comment and specified review person are *required attributes* for a valid EcoSpold1 file. However, in current use many EcoSpold1-datasets are published that have no validation tag at all.¹⁸ This is not ideal, as it is not clear whether these datasets have been reviewed at all or whether merely the specific tag is missing. But it indicates that a missing validation tag does not lead to import errors at least in some LCA software.¹⁹ This means, although formally required, working EcoSpold1-files can under circumstances be lacking the validation tag in actual practice.

If review information is included in XL for a new inventory, the issue arises that a dataset author should not write their own review comments (i.e. the person in 'dataGeneratorAndPublication' should not be the same as the 'proofReadingValidator'). Here, collaboration with the actual reviewer is necessary to combine the reviewer's comments with the dataset author's inventory work. When editing or augmenting an already existing inventory, it should be checked, whether the edits are so large as to deserve a re-review.

While it is encouraged here to perform a proper review with datasets and provide the review information in the appropriate fields, it needs to be recognized that many datasets in current use do not have such information. If Esixell were to *insist* on review information for a proper conversion—and it could as those review attributes are nominally required in the EcoSpold1 format definition—it might lead to a problematic situation for people merely *editing* datasets where no original review information was included. If review information were demanded by Esixell, it could lead to laborious re-review of the edited dataset, when the original datasets was not reviewed—or the original review information was not included or lost during database processing. At worst, this could lead to a charade of a review, where the review information fields are filled with placeholder content, merely to satisfy formal demands, although no *real* independent review has been performed. To avoid this situation, Esixell will allow conversion of files without any of the required review attributes, but if they are not empty, it will convert them.

If a reviewer person is given, Esixell will check against the person defined as datasets generator, and if they are the same will halt conversion. The reviewer shall not be the same person as the author. "Self-review" is at best merely a form of careful internal checking, which is expected anyway, and a proper review must come from an independent, external party.

If a separate reviewer person is given, Esixell will check on the presence of a review comment (i.e. proofReadingDetails, field # 5615). If no review comment is present, although a reviewer person is specified, conversion to ML will be stopped. If a review has actually been performed, the review comment must be given.

5.10 Attribute 'Infrastructure Included'

The attribute 'infrastructureIncluded' is part of the tag <referenceFunction>. In actual practice, it is often not used correctly or as intended by its definition. The actually intended purpose of this attribute is quite special and specific in the case when a dataset is not present as a detailed process inventory

¹⁸ For instance none of the 6801 datasets in the KBOB-UVEK database carry any review or validation information (KBOB et al. 2018). Similarly, in the LCI repository of ESU services at www.esu-services.ch/data/public-lci-reports many of the EcoSpold1 inventories carry no review or validation information. Datasets converted with the old EcoSpoldAccess/Excel2EcoSpold tool seem to have a validation tag.

¹⁹ One such software is SimaPro 9.3 which processed the KBOB 2016 database.

(UPR), but only as a cumulated result from an LCI calculation. I.e. if a dataset consists of some hundreds of biosphere exchanges representing a *cumulated* LCI result, but no technosphere exchanges. Such a cumulated dataset can come for instance from import of confidential datasets from other databases.²⁰ If such a cumulated dataset is being imported, the attribute 'infrastructureIncluded' contains the information, whether the LCI calculation for this cumulated dataset has been performed with or without infrastructure processes included (true/false). Another purpose of this attribute in EcoSpold1 is to indicate in whether the database-wide result calculation were performed with infrastructure included, or with the option to exclude infrastructure. So in cumulated LCI results from the whole database calculation this attribute will be set accordingly. **Thus the attribute 'infrastructureIncluded' is only used and required in cumulated LCI result datasets, not in unit process inventories.**

This actually means that in non-confidential unit processes—which are the much more frequent case—the attribute 'infrastructureIncluded' has *no real purpose*. Also the EcoSpold1 format definition states that in this case the attribute is to remain *empty*. However, the attribute is also formally *required* as per definition, which creates the contradiction that an empty field is formally invalid for this attribute; it *must* be true or false (1/0 in the Excel cell). In past converters, the (formally correct) empty Excel cells were converted into a "false" value in XML.

This somewhat obscure original intended meaning of the attribute—and probably bolstered by the formal incongruity described above—has lead to many misunderstandings in the inventory practice. Some authors consider the field an indication, whether an *unit process inventory* has considered infrastructure burdens or not in its inventory either directly or indirectly, i.e. quite literally whether infrastructure burdens are included in the *UPR*. But also that interpretation is without a point, as perecoinvent methodology *any* infrastructure must be included in unit process inventories anyway, either directly or indirectly.

Esixell will accept cell values of 0/1, as well as empty contents, in the cell for the attribute 'infrastructureIncluded'. **Most likely – unless you actually work with a cumulated inventory of only biosphere exchanges – a value of zero (0 for false) is the appropriate content for this cell.** On Export, empty cells will be overwritten with a "0" (false) value by Esixell, assuring the formally required presence of this attribute and mimicking the behaviour of past converters, thus minimising the risk of incompatibilities with other software and workflows.

Formally likely incorrect is a value of 1 for any UPR, as such an entry would only make sense for *cumulated* LCI results. But Esixell will convert 1 to 'true' nevertheless, as not to possibly disrupt any user's workflow. It is not impossible that a dataset actually represents a cumulated LCI consisting of some thousand biosphere exchanges.

Since in the world of LCA it has become customary to include infrastructure burdens, the whole issue of this obscure attribute can be considered outdated and of little relevance, even disregarding its structural contradictions. However, since it is formally part of EcoSpold1 inventories, it is heeded also in Esixell conversions.

²⁰ The EcoSpold1 format definition states for the attribute 'infrastructureIncluded': "*Indicates whether the unit process imported into the database on the basis of an LCI result (received as cumulative mass- and energy-flows, hence, no LCI results will be calculated for such processes) has included infrastructure processes or not. For all other unit process raw data data sets this data field is empty. After calculation of LCI results in ecoinvent, the data field is filled in according to the fact, whether or not infrastructure has been including during the calculation.*" (ecoinvent 2002).

5.11 Automatic defaults for missing cell contents

Some attributes of EcoSpold1 are formally mandatory to be given by the user. But in several instances, the content of a missing entry is clear or very likely. All those instances are in sheet 'X-Process'. If Esixell encounters following entries to be empty, it will enter a default value and highlight the cell **light green**, for corrected content.

Tab. 5.3 Default values provided by Esixell if missing in the source file.

Mandatory attribute in sheet 'X-Process'	Default value entered when missing	Comment
dataset Type	1	1 stands for <i>unit processes</i> , which is always the case for Esixell application (not exchange definitions or LCIA)
(database) Version	0.00	Since the ecoinvent database version 3+ currently does not use EcoSpold1 anymore, and the former ecoinvent versions 1.00 – 2.20 are closed for new additions or editing as of 2010, no proper ecoinvent version should be given here. 0.00 is used therefore. Inventories elaborated later and merely <i>adhering to the ecoinvent v.2.2 methodology</i> should not be misleadingly presented as having been an actual part of the v2.2 ecoinvent database.
EnergyValues	0	Mandatory, but never used in EcoSpold1. Default is 0.
LanguageCode	en	Usually always 'en' (English)
LocalLanguageCode	de	Usually always 'de' (German)
QualityNetwork	1	Always 1 (originally standing for ecoinvent). Was never extended.
DatasetRelatesToProduct	1	Always 1=true for unit processes (i.e. neither exchanges nor LCIA)
ReferenceFunction Amount	1	Always 1 as per EcoSpold1 definition
InfrastructureIncluded	0	See chapter 5.10 on page 29.
DataValidForEntirePeriod	1	Very commonly this is 1 = yes for all datasets. Flagged as error when missing, but set to 1.
Copyright	1	1 = true. Has no effect on restrictions/confidentiality/openness.

6 Conversion ES1-XML to Excel

6.1 Content in generated Excel

The guide to convert an XML file are described in chapter 'XML to Excel (ML to XL)' on page 9.

The information found in the XML file is written into the appropriate fields of the four sheets of a newly created Excel workbook. Some notable special cases and additions are described below.

6.1.1 Review comments

In chapter 5.9 on page 28 it was explained that the sheet 'X-Process' contains three fields relating to the review process of the datasets and that while these entries are nominally *required fields* according to the EcoSpold1 format definition, in actual practice many EcoSpold1-files lack these entries.

For the conversion from XML files, any present review information will be written out into the Excel workbook, while missing review information will not lead to an error. This reproduces the apparent review status of the XML file. So although formally the review fields are *required* in EcoSpold1, the Esixell Converter will reproduce whatever review information is available from the start file.

6.1.2 ID numbers for sources and persons

According to the EcoSpold1 format definition, each source or person must be given a distinguishing ID number. Sources are usually published reports or papers, but can also be questionnaires or personal communications. The purpose of the ID numbers is that in files with *multiple datasets/activities*, this ID number can be used in sheet 'X-Process' to indicate a particular source or person.²¹ The sheet 'X-Source' then specifies the details of a distinct source (author, year, title etc.) of a given source ID number. The sheet 'X-Person' specifies the details of a particular person (name, address, organisation etc.) of a given person ID number. In an EcoSpold1 Excel file, each source or person must defined centrally only once with their details, and then can be reused as appropriate to indicate sources or persons in various datasets of the same file, without having to duplicate the particular information.

These ID numbers are usually simple integers (1,2,3 etc), but they need to be consistent and without conflict *within one file*, i.e. source 3 should then always mean the *same* source, regardless in which dataset within the workbook this reference is used. In an Excel file, the ID numbers used in 'X-Process' must match the desired entries in 'X-Source' and 'X-Person', especially when several datasets are in one file. There is however no formal requirement for universally consistent ID numbers *per project* or *all EcoSpold1 files*. So source ID 3 can mean a different publication in *different* XML files.

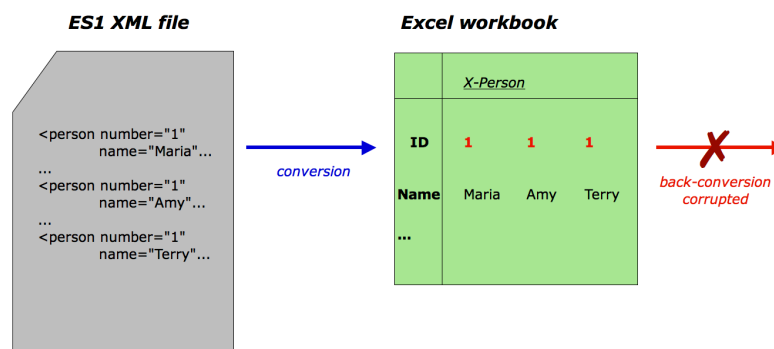
Problems with XML files written by third parties

When an Excel file is converted into an XML file, the particular information contained in 'X-Source' and 'X-Person' is *duplicated* and written into each dataset tag, along with the ID number that was used. If an XML file contains *multiple* datasets, these ID numbers should therefore be different for different

²¹ In the sheet 'X-Process', *sources* are indicated in the attribute 'ReferenceToPublishedSource' of the tag <dataGeneratorAndPublication>, while *persons* are indicated in three different places: in the tag <dataGeneratorAndPublication>, i.e. the person who compiled the inventory figures; in the tag <dataEntryBy>, i.e. the person feeding the database; and in the attribute 'proofReadingValidator' of the tag <validation>, i.e. the person performing the review. So even within one single dataset up to three different persons need to be properly distinguished, and identical persons must be addressed with the same ID number.

sources/persons. Unfortunately, **some XML files in use do violate the EcoSpold1 format definition** by giving *each and any* source the same ID number "1".²² When converting such a non-conform XML file back into an Excel table as done with Esixell, the problem arises that *different* sources have been given an *identical* ID number. If in 'X-Source' all entries would feature the ID number "1", the number would become useless, since that number will not unambiguously specify the desired source in the Excel tables anymore. While the source XML file (with duplicated source and person information within *each* dataset entry) is unambiguous and apparently works for the intended LCA software, without any adjustments their back-conversion from Excel to XML (cf. Fig. 1.2) could produce an ambiguous or even faulty file which is incompatible with the EcoSpold1 format definition.

The following image illustrates the problem with indiscriminating ID numbers given for all persons.



Solutions for the Esixell Converter

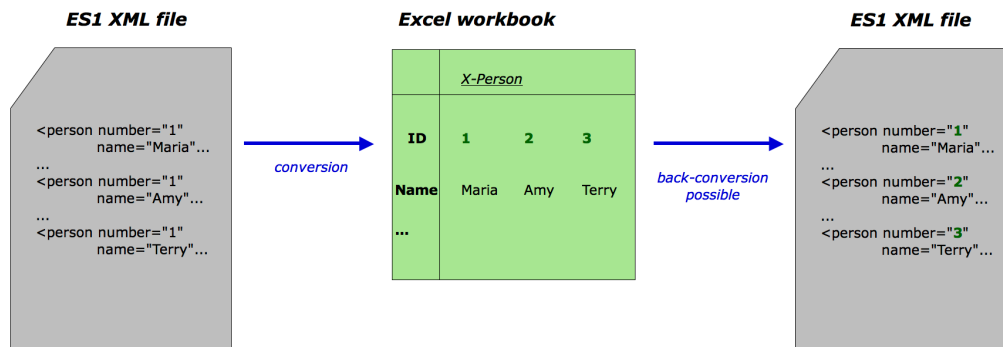
Esixell will automatically try to convert the erroneous ID numbers into a new set of consistent, unique IDs in the Excel tables. It will try to keep any original ID numbers that are not in conflict.

This issue of useless ID numbers in start XML files can be solved by disregarding identical ID numbers for apparently different entries, and replacing one of the ID numbers with a new and actually unique ID number. Again: the ID numbers must only be unique within one file. They don't need to be unique and consistent within a project or all EcoSpold1 files (but they can be, of course). There is no requirement that used ID numbers in a file must be consecutive or be starting from 1. Esixell will therefore simply look at the parsed ID numbers up to that point and use the largest value plus 1 to derive a new replacement ID number.²³

The following image illustrates the solution of overwriting indiscriminating ID numbers for persons with new, properly discriminating ID numbers.

²² The KBOB 2016 database, which is published as one large EcoSpold1 XML file containing 6800 datasets, has this problem, as all of its 129 different source entries all feature the ID number "1", which in this instance has lost its usefulness to identify anything.

²³ This straightforward approach can in theory lead to valid ID numbers being reassigned nonetheless. For instance if ID number 8 was used entirely correctly in the original file, it is still possible for it to be replaced, if *previous* datasets in that same file have already made it necessary to create new ID numbers of 8 and larger. This problem seems of little practical relevance, as hardly any authors do keep a list of unique ID numbers, whose workflows could be disrupted by this automatic re-assignment.



Distinguishing referenced sources

In the EcoSpold1 format definition, the first author's name, publication year, and publication title are required fields to specify a source.²⁴ Sources, where these three fields have the same content, can be considered identical. So instead of identifying sources by their (possibly unhelpful) ID number, they can reasonably be identified by their combined content of author+year+title. When converting an XML file to an Excel table, this content info for author+year+title can be used to decide if a new, different source is being defined. If so, this new source can be written as an entry in 'X-Source'. To produce a valid and conform EcoSpold1 file, in cases where XML files contain the same ID numbers for different sources, the ID number must be *replaced* by a useful, actually unique ID number. This new ID number is then also used in 'X-Source' (in the field 'number' #801) and 'X-Process' (in the field 'ReferenceToPublishedSource' #757). Such a harmonised ID number achieves that the produced Excel table conforms with the EcoSpold1 format and can also be properly *converted back* into a valid XML file, if so desired.

Distinguishing referenced persons

A similar solution can be applied for inadequate ID numbers for person entries.²⁵ In EcoSpold1 format definition, a person's name is a required field, along with their address, telephone number, company affiliation, and country. It can be useful to distinguish people's employer with the company information, because different projects at different times could be executed under different employment even by the same person. So to distinguish different person entries, the field 'name' and 'companyCode' are considered. Entries, where the entity name+company is the same content, can be considered identical. Also here, if the original ID numbers in the start file are non-unique, they need to be overwritten by a useful, actually unique ID number. This new ID number is then also used in 'X-Person' (in the field 'number' #580) and 'X-Process' (in the field 'dataGeneratorAndPublication' #751 or 'dataEntryBy' #302 or 'proofReadingValidator' #5616).

²⁴ Additionally, also type of source (separate publication, oral communication, questionnaires etc.) place of publication (location and/or web address) are required fields. But the other three fields are likely sufficient to sensibly distinguish different publications.

²⁵ The XML file of the KBOB 2016 database has always given the ID number "3" to the data generator (author) and the ID number "2" to the data entry person, even when they were the *same* person.

Repository of information

Since the above procedures might change information contained in the start XML file, the original information on referenced sources and persons of each dataset/activity, including any original, inadequate ID numbers, is reproduced in the Excel tables at the bottom rows of 'X-Process' in blue font colour. These cells are informal only and do not have any bearing at all on the back conversion XL to ML. Converted is only the uninterrupted, consecutive range of cells at the top (cf. chapter 5.5 on page 24).

The repository of information at the bottom of 'X-Process' is informal only, is separated by the formal part by an empty row, and its contents will *not* be back-converted into XML. **The blue font colour indicates to the user, that this content is informal only.**

An example of such an informal information repository is shown in Fig. 6.1 below.

Remaining problems

Esixell cannot remedy all problems concerning consistent lists of sources and persons. Problems arise for instance when identical sources are referenced with different defining content fields. As an example, following two entries for sources were found in an XML file:

- Entry 1: *Jungbluth, N., Stucki M., Fluri K., Frischknecht R., Buesser S., (2012) 11Life Cycle Inventories of Photovoltaics. Ed. Bauer C., by ESU-services ltd., ESU-services ltd., Vol.407, www.esu-services.ch.*
- Entry 2: *Jungbluth, N., Stucki M., Fluri K., Frischknecht R., Buesser S., (2012) 1Life Cycle Inventories of Photovoltaics. Ed. Bauer C., by ESU-services ltd., ESU-services ltd., Vol.407, www.esu-services.ch.*

These two entries are obviously for the same source publication, but were recorded with slightly different publication *titles*, i.e. with a leading "1" or a leading "11", formally making these two into two *different* sources. Esixell is not smart—or bold—enough to detect or decide that these two sources are actually the same. The users must scan for, inspect and remedy such errors themselves.

Another example encountered is from the X-Person entries for the name of person attribute:

- Entry 1: *name="Thomas Kägi"*
- Entry 2: *name="Thomas Kägi, Telephone: 0041 61 206 95 58; E-mail: ecoinvent@carbotech.ch; Company: Carbotech AG; Country: CH"*

Also here these two formally distinguished entries are for the same person. In the second instance, the person's name was combined with all contact information and written as one long string into what is supposed to be only the person's *name* field, while this contact information is also repeated in their proper respective attribute fields.

Esixell cannot correct or detect either of these errors. It is the author's task to avoid such unnecessary duplications by creating a consistent and carefully maintained list of sensible entries (cf. also footnote 8 on page 24).

Dataset	200	number	1
	none	generator	SimaPro 8.3.0.0
	none	timestamp	2014-11-17T16:21:45
		[name of original XML file]	Top 237 ES1 KBOBv2_2016.XML
		[error messages]	<ul style="list-style-type: none"> • Required field dataValidForEntirePeriod was undefined. • Timeperiod of dataset seems wrong.
Sources	757	original source ID number	1
		used ID number	1
		SourceType	4
		Text	CD-ROM
		FirstAuthor	Nemecek, T.
		AdditionalAuthors	Kägi, T., Blaser, S.
		Year	2007
		Title	Life Cycle Inventories of Agricultural Production Systems
		PageNumbers	
		NameOfEditors	
		TitleOfAnthology	Ecoinvent report version 2.0
		PlaceOfPublication	Swiss Centre for LCI, ART
		Publisher	Swiss Centre for LCI, ART
		Journal	
		VolumeNo	15
		IssueNo	
Data Entry person	302	original person ID number	2
		used ID number	2
		Name	Jürgen Sutter
		Address	Unknown
		Telephone	0041 44 633 44 73
		Telefax	
		Email	eth.s-u@ecoinvent.org
		CompanyCode	ETH S+U
		CountryCode	CH
Data Generator Person	751	original person ID number	3
		used ID number	2
		Name	Jürgen Sutter
		Address	Unknown
		Telephone	0041 44 633 44 73
		Telefax	
		Email	eth.s-u@ecoinvent.org
		CompanyCode	ETH S+U
		CountryCode	CH
Validator Person (Review)	5616	original ID number	
		converted ID number	
		Name	
		Address	
		Telephone	
		Telefax	
		Email	
		CompanyCode	
		CountryCode	

Fig. 6.1 Example of the informal information repository on the bottom of sheet 'X-Process'. The blue font colour shall call attention to the fact that any editing of information here will *not* be relevant for any back-conversion XL to ML. It is separated from the proper inventory by an empty row.

6.1.3 Dataset number, generator and timestamp

EcoSpold1-XML-files contain three attributes which are required according to the EcoSpold01 format definition, but are entered *automatically* by the processing software.

The "number" attribute of the <dataset> tag is simply a sequential count number for each dataset that is in a given XML file, starting from 1. In files with several inventories this represents an ID number of an activity, but it is never used anywhere to reference an activity.

The "generator" attribute contains a text string specifying the *processing software*, and in existing XML-files strings like e.g. "SimaPro 7.3" or "EcoInvent Excel2EcoSpold 1.9.17" (from the old EcoSpoldAccess Add-in) are found here.

The third attribute is "timestamp" which contains a date and time of when that inventory was processed by the software, e.g. "2010-06-22T16:32:39".

For each process inventory, either those attributes can be present in EcoSpold1-XML in the tag <dataset>. But neither of those attributes can be manually assigned by the dataset author, as they are entered automatically. When Esixell converts XML files to Excel workbooks, for purely informational purposes any found dataset number, generator name or timestamp information is displayed in 'X-Process' in a separated *informal repository* at the bottom of the sheet in blue font colour (cf. Fig. 6.1 on page 36). For transparency during manual editing of datasets this information might be of assistance.

Please note that when converting such an Excel file back to an XML-file, those informal cells will *not* be read, but instead the dataset numbers will be given according to the sequence encountered in the start file. The timestamp will be automatically updated to the current time of conversion. The generator string will be set to "Esixell Converter v1 – 2023 ND".²⁶ It seems inappropriate to leave an old, former generator name, when the actual XML file has not really been generated by that previous software.

6.1.4 Internal Version Number

The attribute `internalVersion` of the tag <dataSetInformation> nominally contains a version number of a dataset and according to the definition "is used to discern different versions during the working period until the dataset is entered into the database". So this is purely an internal housekeeping attribute of the ecoinvent database during the review process. It is automatic and cannot be set by the author of the dataset. In actual practice the internal version number frequently reads "0.00" or "1.00" and does not contain any tangible dataset version information. When writing XML files, Esixell will therefore use "1.00" as `internalVersion`. When reading XML files any data in `internalVersion` is not parsed.

6.1.5 Start XML file name

For datasets converted from a start XML file into an Excel workbook, it might be helpful for transparency, if the originating file is indicated. The name of the converted XML file is shown in the informal repository at the bottom rows of the sheet 'X-Process' in blue font colour (cf. Fig. 6.1 on page 36).²⁷ That field has purely informational purposes and is not part of EcoSpold1. Please note that when converting such an Excel file back to an XML-file, that field will therefore *not* be read. There is no tag or attribute in the EcoSpold1 format designed to contain the originating file, and in the case of XL to ML that name would be the name of the start *Excel file* anyway.

If a dataset author desires to include the name of the originating Excel file into an EcoSpold1 XML file, it could be included for instance in the `GeneralComment` field or possibly in the `proofReadingDetails`.²⁸

²⁶ The last two letters indicate the user's currently chosen preference on handling of duplicate exchanges: "ND" means "no duplicates allowed"; "DA" means "duplicates allowed". See setting preferences in chapter 3.1.1 'Allow duplicate Exchanges?' on page 8.

²⁷ Also, Esixell will write the name of the original file into the Excel document properties of the generated workbook as a comment.

²⁸ Instead of manually entering the filename into the general comment field, Excel has a worksheet function that flexibly produces the file name of a workbook =CELL("filename", A1). Here using the cell address A1 as an arbitrary cell of the intended source workbook. In German versions of Excel it's =ZELLE("Dateiname", A1). This function contains the *entire*

6.1.6 Column layout in X-Exchange

The generated sheet 'X-Exchange' contains the parsed biosphere and technosphere exchanges for each dataset. The column layout is initially as shown in Tab. 5.2 on page 18. Thus per dataset up to 7 columns are used in 'X-Exchange'. However the columns for MinValue, MaxValue, and MostLikelyValue are only sensible for triangular and uniform distributions. As these distributions are rarely used in practice, most of the time, these columns are superfluous and only take up space in the workbook. After a parse to XL, Esixell will therefore delete those columns, if none of the listed exchanges of the dataset requires them, for compactness and reduced file size. So oftentimes a dataset will only occupy 4, not 7, columns (meanValue, UncertaintyType, StandardDeviation95%, GeneralComment).²⁹

6.1.7 Text formatting in column for Exchanges General Comment

The last column for each process is the General Comment for each exchange. These can be lengthy text entries giving details to each exchange. With Excel's standard cell format, long text can spill over to the right and make reading of other datasets cumbersome, as shown in the following image.

file path with folder hierarchies. To obtain the file name alone without path information, the string can be trimmed with following formula, here assuming that cell D160 contains the above function for the file path: =MID(D160,FIND("[",D160)+1,FIND("]",D160)-FIND("[",D160)-1). Or in German Excel versions: =TEIL(D160;FINDEN("[";D160)+1;FINDEN("];D160)-FINDEN("[";D160)-1). Excel's string operations can be used to combine a GeneralComment content from fixed and flexible text parts, like = "Inventory refers to the production of 1 kg ... etc." & "Originating file name is " & D161. Here assuming cell D161 contains the file name function outlined before. Such a string formula can be placed directly in the appropriate cell for the general comment in sheet 'X-Process'. Esixell will read out resulting cell values for conversion, regardless, whether the cell contains a formula or not.

²⁹ GeneralComment is not formally a required field in EcoSpold1, but it is abundantly used—and highly recommended—, for explaining and documenting exchanges, as well as Pedigree matrix information. Its column is therefore never deleted in Esixell, even when entirely empty.

3707	3708	3709	3792	3707	3708	3709	3792
feldspar, at plant	Uncert aintyT ype	StandardD eviation95 %	GeneralComment	disposal, municipal solid waste, 22.9% water, to sanitary landfill	Uncert aintyT ype	StandardD eviation95 %	GeneralComment
RER 0 kg				CH 0 kg			
8.4307E-14	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
9.31E-16	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
8.09E-16	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
6.11E-16	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
1.8383E-13	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
3.16E-16	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
0	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
1.977E-15	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
5.65E-16	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
1.5121E-09	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
1.0986E-14	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
1.5046E-09	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
4.9897E-08	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
3.0788E-09	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
3.7363E-11	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
3.9125E-12	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
4.8777E-11	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
9.4346E-09	0		burden from incineration of slud	0.00000145	1	559.0%	Emissions from short-term lead
1.5344E-08	0		Emissions from short-term lead	0.00000141	1	329.0%	Emissions from short-term lead
3.8891E-12	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
1.146E-14	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
4.12E-16	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
2.8566E-12	0		Emissions from short-term lead	6.36E-12	1	1990.0%	Emissions from short-term lead
1.0731E-10	0		Emissions from short-term lead	8.46E-12	1	1530.0%	Emissions from short-term lead
4.5E-17	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
1.1401E-10	0		Emissions from short-term lead	1.03E-08	1	813.0%	Emissions from short-term lead
1.9494E-11	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
1.3694E-07	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
1.1152E-14	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
4.88E-16	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
4.9646E-09	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
9.623E-14	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
2.4155E-13	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
3.4898E-12	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
1.3283E-12	0		Emissions from short-term leachate treatment and incineration of resulting sludge. Uncertainty calculated depending on waste input and				
4.3888E-10	0		burden from direct release or in	0.000000425	1	1010.0%	Emissions from short-term lead
0	0		burden from direct release or incineration of landfill biogas. Uncertainty calculated from uncertainty of the waste composition of this waste,				
2.6559E-10	0		burden from direct release or incineration of landfill biogas. Uncertainty calculated from uncertainty of the waste composition of this waste,				
5.21E-16	0		burden from direct release or incineration of landfill biogas. Uncertainty calculated from uncertainty of the waste composition of this waste,				
4.1613E-07	0		burden from direct release or incineration of landfill biogas. Uncertainty calculated from uncertainty of the waste composition of this waste,				
4.8509E-09	0		burden from direct release or incineration of landfill biogas. Uncertainty calculated from uncertainty of the waste composition of this waste,				
2.3476E-14	0		burden from direct release or incineration of landfill biogas. Uncertainty calculated from uncertainty of the waste composition of this waste,				
1.3407E-10	0		burden from direct release or in	3.81E-12	1	1300.0%	Emissions from short-term lead

Fig. 6.2 Long texts in Exchange's GeneralComment could spill over to the right in Excel's standard cell format.

A solution here would be to allow wrapping of texts into multiple lines. But this makes the row height and the table layout irregular. Esixell will instead format the Exchange's GeneralComment cells with a so-called *horizontal fill alignment*, where the text is used to fill out the cell, but not beyond the cell border, resulting in a uniform, clean and legible table, without impeding other datasets.

3707	3708	3709	3792	3707	3708	3709	3792
feldspar, at plant	Uncert aintyT ype	StandardD eviation95 %	GeneralComment	disposal, municipal solid waste, 22.9% water, to sanitary landfill	Uncert aintyT ype	StandardD eviation95 %	GeneralComment
RER 0 kg				CH 0 kg			
8.4307E-14	0		Emissions from short-term lead				
9.31E-16	0		Emissions from short-term lead				
8.09E-16	0		Emissions from short-term lead				
6.11E-16	0		Emissions from short-term lead				
1.8383E-13	0		Emissions from short-term lead				
3.16E-16	0		Emissions from short-term lead				
0	0		Emissions from short-term lead				
1.977E-15	0		Emissions from short-term lead				
5.65E-16	0		Emissions from short-term lead				
1.5121E-09	0		Emissions from short-term lead				
1.0986E-14	0		Emissions from short-term lead				
1.5046E-09	0		Emissions from short-term lead				
4.9897E-08	0		Emissions from short-term lead				
3.0788E-09	0		Emissions from short-term lead				
3.7363E-11	0		Emissions from short-term lead				
3.9125E-12	0		Emissions from short-term lead				
4.8777E-11	0		Emissions from short-term lead				
9.4346E-09	0		burden from incineration of slud	0.00000145	1	559.0%	Emissions from short-term lead
1.5344E-08	0		Emissions from short-term lead	0.00000141	1	329.0%	Emissions from short-term lead
3.8891E-12	0		Emissions from short-term lead				
1.146E-14	0		Emissions from short-term lead				
4.12E-16	0		Emissions from short-term lead				
2.8566E-12	0		Emissions from short-term lead	6.36E-12	1	1990.0%	Emissions from short-term lead
1.0731E-10	0		Emissions from short-term lead	8.46E-12	1	1530.0%	Emissions from short-term lead
4.5E-17	0		Emissions from short-term lead				
1.1401E-10	0		Emissions from short-term lead	1.03E-08	1	813.0%	Emissions from short-term lead
1.9494E-11	0		Emissions from short-term lead				
1.3694E-07	0		Emissions from short-term lead				
1.1152E-14	0		Emissions from short-term lead				
4.88E-16	0		Emissions from short-term lead				
4.9646E-09	0		Emissions from short-term lead				
9.623E-14	0		Emissions from short-term lead				
2.4155E-13	0		Emissions from short-term lead				
3.4898E-12	0		Emissions from short-term lead				
1.3283E-12	0		Emissions from short-term lead				
4.3888E-10	0		burden from direct release or in	0.000000425	1	1010.0%	Emissions from short-term lead
0	0		burden from direct release or in				
2.6559E-10	0		burden from direct release or in				
5.21E-16	0		burden from direct release or in				
4.1613E-07	0		burden from direct release or in				
4.8509E-09	0		burden from direct release or in				
2.3476E-14	0		burden from direct release or in				
1.3407E-10	0		burden from direct release or in	3.81E-12	1	1300.0%	Emissions from short-term lead

Fig. 6.3 Exchange's GeneralComment cells with horizontal fill alignment prevent spilling over of text onto empty cells on the right.

A minor inconvenience of this solution is that when the GeneralComment is *shorter* than the column width, the text might be *repeated* within the cell display. E.g. if the entire comment is but a short text value like "UVEK", it might be displayed as "UVEKUVEKUVEKUVEKUVEK". This is not an error, and the cell's *value* is correct and will be back-converted by Esixell as "UVEK" only, but the *display* of the cell will show a repeated content. The user is completely free to change this cell formatting in Excel, as this will not affect Esixell's parsing. But such short comments are ill-advised anyway and it is suggested to expand them to give more information for the exchange concerned.

6.2 Warnings and errors from ML to XL

The start XML files for conversion must adhere to certain formal requirements. The Esixell converter cannot fully validate an XML file, e.g. for typos or for canonical existence of certain exchanges. It will not check for instance whether 'DF' is a proper country code or a typo, and it will not check for instance if the exchange 'dimethyl mercury//air/high population density//kg' exists canonically or not. But Esixell tries to detect several other possible inconsistencies or errors. In some cases it can correct and remedy the inconsistencies by itself. In other cases, it alerts the user/author to those inconsistencies, which must then be remedied by the user.

Cells with an orange fill colour contain questionable content which must be controlled by the user and usually corrected for a valid Excel-EcoSpold1 file to result.

A full summary text of any encountered errors is given for each dataset in sheets 'X-Process' in row 54 (informal repository). If errors are encountered, the cell with this **error summary text** is selected and displayed in Excel (or the last dataset with errors in case of multiple datasets per workbook). Conversely that means if after an import the workbook remains on the first sheet "X-Exchange" that Esixell has not detected any errors in the source XML file.

Cells with a **light green colour** have been automatically corrected by Esixell during the parsing and do not require correction (but the highlight indicates something was altered compared to the start XML file). Errors with an **orange cell fill** must be checked by the user and usually corrected.

StartDate	1900-01-01
------------------	------------

Fig. 6.4 Example of Esixell highlighting **questionable content** the user needs to check/remedy.

ReferenceToPublishedSource	2
----------------------------	---

Fig. 6.5 Example of Esixell highlighting **content corrected** automatically for consistency.

Errors can be detected and highlighted in all four worksheets. The user can use the error summary text to locate and correct all errors. Usually also the offending highlighted cell receives a cell comment mentioning the error.

In the following chapters some types of errors are explained. There are formal errors and content errors, which are distinguished with a prefix in the chapter's title:

[] = formal errors

\$ = content errors

6.2.1 [] Unique exchange entries

In an inventory for a process/activity, each of the exchanges listed should only be used once. A particular inventory figure represents always a distinct location in a matrix.³⁰ Each location in a matrix—or matrix element—can only be occupied by a single inventory value. If more than one inventory value exists for a particular matrix element, an ambiguity is created. Shall the matrix element contain a *summation* of the multiple occurrences? What shall happen to the exchange text comments in the inventory? Are any of the multiples an actual error—and if so, which one is to be removed and which one is to be kept?

Resolving duplicate exchange information

The user can set preferences in Esixell on how encountered duplicate exchanges shall be treated (see chapter 3.1.1 'Allow duplicate Exchanges?' on page 11). If the user—against recommendation—allows duplicate exchanges, the duplicates will be written into the formal part of the inventory list in X-Exchanges. On back-conversion XL→ML those duplicates will be included.

If duplicate exchanges are not allowed by the user and Esixell detects duplicate exchanges in the same dataset/activity, it will highlight the table cell as questionable content (cf. Fig. 6.4), add a cell comment, but it will not overwrite the existing inventory content from the first instance, in order not to destroy any original information. Thus, any subsequent repetitions of exchange entries within an inventory are ignored formally.³¹ But the information of the duplicate exchange (mean value, uncertainty information and exchange comment) is written into an *informal repository* of duplicate exchanges at the bottom rows of sheets 'X-Exchange' in blue font.³² It is then the task of the user/author to remedy the detected ambiguity and decide:

- a) if the duplicate entry was in error and should either be removed or renamed, or
- b) if the first entry was erroneous and should be removed or renamed, or
- c) if the duplicate entry is occurring in the real world and therefore should be joined (added) to the first instance of the exchange.

New or additional exchanges from this procedure need to be included in the *formal* upper part of 'X-Exchange'. Any changes in the *informal* repository at the bottom (blue font) will be ignored in back conversion into XML.

6.2.2 [] Unique activity entries

Within a file—and indeed within a database—a particular activity can only occur once. Defining attributes to distinguish activities/datasets are the dataset name, the geographic location, the boolean for infrastructure, and the physical unit. These are the first four entries of each dataset column in 'X-

³⁰ The matrix usually named A contains all links between the technosphere processes, so is a square matrix with all technosphere processes on each dimension. The matrix usually named B contains the links between the technosphere processes and the biosphere, the elementary exchanges.

³¹ I.e if an Excel table created from XML by Esixell were to be converted back into an XML file, only the first exchange instance would be converted, but not the subsequent duplicates.

³² The duplicates are informally listed in the sequence they are encountered and each duplicate gets a separate line. No effort is made to harmonize duplicates on different lines or amongst different datasets. The reason for this is that in encountered ES1-XML also files with *triplicate entries* were encountered. If the duplicate list were organised into containing only one line per unique exchange, it could overwrite exchange information of previous duplicates.

Exchange' and 'X-Process'. If two activities have the same content in all of those four attributes, they are nominally identical. Certain LCA software might not be accepting duplicate activities.³³

Duplicate datasets are questionable. They might come from typos or copy-paste errors. If in a workbook several identical processes occur, Esixell will highlight the header for duplicate processes in 'X-Exchange' and 'X-Process' in orange. The user must decide, if the duplicate or the first instance are to be deleted, or if other changes to harmonize the datasets are necessary.

If each activity is parsed into a *separate* workbook, no checking for duplicates will be performed by Esixell.

6.2.3 [] Unique ID numbers for sources and persons

As outlined in chapter 6.1.2, literature sources and persons must be distinguished in valid XML files with consistent ID numbers. When converting from ML to XL, Esixell will consolidate the found entries. This can mean overwriting ID numbers found in the original XML.³⁴ When Esixell overwrites ID numbers of sources and persons in order to create consistent entries, it will colour the respective Excel cells with a **green fill colour**, to highlight the corrected content (cf. Fig. 6.5). Changed ID numbers will also be highlighted in green in sheets 'X-Source' and 'X-Person'.

These changes can influence the further work of the user. If for instance a user is accustomed to a certain ID number representing a certain source or person, Esixell might have changed this number for internal consistency.

6.2.4 \$ Time period dates for year 1900

In EcoSpold1 files produced from the SimaPro LCA software, the date fields for startDate and endDate usually read both "1900-01-01". This is apparently a data handling error, possibly not recognizing the specified date text string. In datasets referring to the year 1900 as a start date, a warning comment for the respective cell is created and the start date cell is highlighted orange.

6.2.5 \$ Date format errors

The expected date formats for startDate and endDate are text strings with the sequence 'yyyy-mm-dd', e.g. '2022-01-31'. If a source XML file contains strings not adhering to this format, Esixell will create a warning comment and colour the cell orange. Esixell will not try to guess the correct date from an ill-formatted string.³⁵

³³ For instance in OpenLCA the import routine for EcoSpold1 called 'EcoSpold01Import.java' will refuse to import activities into a database that already exist.

³⁴ For instance, if person A is the person for DataEntryBy and also for DataGeneration, but in the original file different ID numbers are given for this same person A, the ID numbers must be consolidated by Esixell to point to the same entry in X-Person. Similarly if two different persons were given the same ID number in the original file, the ID numbers must be differentiated to be sensible and formally valid. The identity of a person is derived in Esixell by looking at the combination of 'person name + company name'. The identity of a source is derived by Esixell from the combination of 'first author name+year+title.'

³⁵ However Esixell will convert strings like 'yyyy-mm' or 'yyyy' into full dates. See chapter 5.7 on page 25.

6.2.6 \$ Incompatible time period dates

Esixell will check if the given start date is *not later* than the given end date. For erroneous sets of dates, a warning comment for the end date cell is created and the cell is highlighted orange. EndDates being equal to the StartDate, even up to the day, are accepted, even if a day-specific inventory is very unusual.

6.2.7 \$ Empty strings for required content

Some fields in EcoSpold1 are mandatory. In some cases, their content is clear. For instance the attribute 'energyValues' is formally required, but always 0, as it was never operationalised. Or the attribute 'qualityNetwork' has never any other value than 1, standing for ecoinvent. See overview in Tab. 5.3 on page 31.

For some missing mandatory attributes their content is not constant, most evidently in a mandatory text field like the title of a publication source, which shall not be empty. Esixell will highlight cells, where expected content is missing or questionable.

In Ecospold1 XML exports from the SimaPro LCA software the field 'dataValidForEntirePeriod' is often given merely as an empty string, i.e. `dataValidForEntirePeriod=""`. A correct entry for this mandatory field in XML would either be `dataValidForEntirePeriod="true"` or `dataValidForEntirePeriod="false"`. It is not clear what an empty string signifies and this is possibly the result of a past data handling error.

For most datasets "true" will be the correct choice here. Esixell will interpret any empty strings here as true (=1). Since this might be a questionable guess for a particular dataset, the cell for this attribute in 'X-Process' will receive an error comment and an orange fill colour, for the user to review and possibly correct.

6.2.8 \$ CAS numbers inconsistent checksum digit

CAS numbers are codes to identify certain chemicals or compounds. A CAS number has the following format "000000-00-0".³⁶ The last digit is a checksum digit calculated from the preceding digits. When Esixell encounters a CAS number (with or without any leading zeros) it verifies this checksum digit. When inconsistencies are found, the cell with the CAS number in sheet 'X-Process' is highlighted in orange and a cell comment is added, alerting to the wrong CAS number. The user must then research the appropriate CAS number. It might be that for the intended chemical the checksum digit is a typo or that there is a typo in the preceding digits, and Esixell cannot decide this. The user must then overwrite the appropriate Excel cell with the correct CAS number for the compound/material.

6.2.9 \$ Uniform and triangular uncertainty distributions

Though rarely used, it is possible to have a triangular or uniform uncertainty distributions for exchanges. The uncertaintyType = 3 is for triangular distribution; and = 4 for uniform distributions. These distributions are not described with the field 'StandardDeviation95%' but with the fields 'MinValue' and 'MaxValue'. Esixell checks if the given mean value is in the range of the given

³⁶ Any leading zeros in a particular CAS number could be omitted in EcoSpold1.

minimal and maximal value, and also whether the minimal value is lower or equal to the maximal value.

For the triangular distribution also the given mean value must be located in the *middle third* of the range spanned by minimal and maximal value. The given mean value is *not* the "tip" of the triangular distribution. The tip of the triangle is the maximum likelihood value (MLV), while the mean is the arithmetic mean of the entire distribution, which is equivalent to the position of the centre of gravity, if the triangle were a solid matter triangle.³⁷ Unless $\text{min} = \text{max} = \text{MLV}$, i.e. a collapsed distribution without any uncertainty, the mean value can never be at the extremes of the distribution (only the MLV can), and the centre of gravity must be within the middle third of the min–max range. If Esixell encounters any incompatibilities for triangular distributions regarding this, it will highlight an error.

The maximum likelihood value (MLV) is formally not used in EcoSpold1, as it can be inferred from mean, min and max.³⁸ If however an XML file specifies a MLV it will be displayed in the appropriate column. Although formally ignored during back conversion, an MLV might help in deciphering and correcting errors for a given triangular distribution.

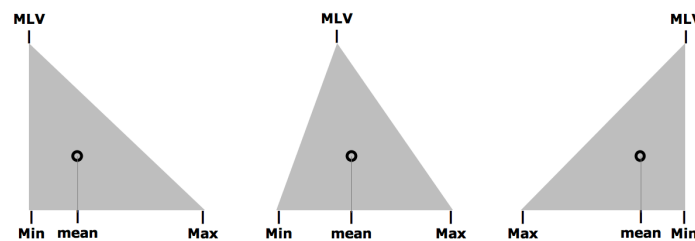


Fig. 6.6 Schematic examples of triangular distributions and the location of the maximum likelihood value (with label "MLV") and the arithmetic mean (with marker \circ).

6.2.10 Uncertainty info given for the reference product exchange

In the list of exchanges, each exchange must have at least an `UncertaintyType`, *except* for the reference product exchanges or by-products. The `UncertaintyType` defines the type of uncertainty distribution (cf. chapter 10.1 on page 53). Formally and officially, the reference product exchange must not have an `UncertaintyType` attribute, i.e. it is missing entirely.

In some inofficial, mutated 'dialects' of EcoSpold1 this rule might not be adhered to. In recent (2023) versions of ES1 files exported from SimaPro, the reference products were given an `UncertaintyType` of zero ("0" for "undefined"), when previously SimaPro adhered to the official conventions. Some beta testers even reported that SimaPro now even *requires* such an entry to be present in imports of ES1 files, or else denies import with an "attribute missing" error. **Thus depending on the target software an non-empty entry of `UncertaintyType` for the reference product exchange is either incorrect (official convention) or required (occasionally encountered in recent SimaPro).**

If Esixell parses an XML file into an Excel table, it will highlight either entries for `UncertaintyType` or `StandardDeviation95%` for the reference product exchange, if they are not empty. It will comment in

³⁷ The mean value can be calculated from $(\text{min} + \text{max} + \text{MLV}) / 3$. The MLV can in turn be calculated from $(3 \cdot \text{mean} - \text{min} - \text{max})$, and must be $\geq \text{min}$ and $\leq \text{max}$.

³⁸ The EcoSpold01 schema definition defines a maximum likelihood value (called "mostLikelyValue", ID 3797), but somewhat paradoxically rejects its use already within the schema definition, in favour of defining the mean value instead.

the error messages that this might be OK for files exported to SimaPro (in which case probably no correction is required).

Upon re-export as XL → ML, Esixell will include whatever uncertainty attributes are given for the reference product, also `UncertaintyType = "0"`, as this might be suitable for SimaPro re-import. **Therefore the user has full control over the produced XL output, as needed.** For `uncertaintyType = "0"`, i.e. undefined, Esixell will ignore any figures given in the column for `StandardDeviation95%` upon re-export XL → ML. For uncertainty types other than "0" the usual requirements apply, also for the reference product: if a lognormal distribution is indicated (`UncertaintyType = "1"`) then the `StandardDeviation95%` cannot be empty etc.

6.3 Tips for the start XML file

The start XML files usually come from LCA software and not much influence on content or formatting can be exerted by the user. XML files are commonly simple text files with the file suffix `.xml`.


6.3.1 Delimiter for new line

One aspect though can potentially be changed by subsequent handling of files, which is the characters used to indicate a new line, the line delimiter. This is relevant since Esixell reads and parses the XML files in a *line-wise fashion* and each XML tag is expected to be written onto a separate line. Unfortunately even in simple text files, new line delimiters are not uniform and can depend on the originating computer operating system.

		New line delimiter
Microsoft Windows text files	DOS	CR + LF
Apple Macintosh text files	Mac	CR
Unix text files	Unix	LF

where *CR* = Carriage return character, Unicode 13
LF = Line feed character, Unicode 10

Most EcoSpold1 XML files adhere to the DOS text file convention (CR+LF). Esixell conversion seems to work well with either DOS or Mac files, but with Unix files the *whole* text document is read as one long, single line, which cannot be parsed properly in Esixell.

If for some reason your start XML file is a Unix text file, you need to change it into a DOS or Mac text file to use it with Esixell. For this you can use free text editors.³⁹ For instance in editors like TextWrangler or BBEdit, choosing the toolbar icon 'File Options' () allows you to switch the text file type to the required DOS and upon a resave the kind of file will be changed.

³⁹ A list of text editors is given in Wikipedia, showing which ones support the feature 'Newline conversion' https://en.wikipedia.org/wiki/Comparison_of_text_editors#Basic_features. Instead of such desktop software, also several command line conversion commands can be used.

6.3.2 File size and maximal number of datasets

The conversion ML to XL is limited by the number of Excel table columns that can accommodate dataset entries in sheet 'X-Exchange'. A separate dataset requires at least 4 columns in 'X-Exchange' (up to 7 columns when MinValue, MaxValue, MLValue are used). Since Excel sheets can have up to 16'384 columns, there is space for up to 4090 datasets.

There can be advantages of having that many datasets in one big table, but for editorial work, this many datasets might be detrimental. Esixell accepts any size of XML files for conversion, provided the <dataset> tags are provided in full and not cut apart. Thus it is possible to manually divide large files into smaller sections in a simple text editor at a juncture </dataset><dataset>. For use with Esixell as start files, these separate sections need not to be complete ES1-XML files with header information tags like '<ecoSpold xmlns="http://www.EcoInvent.org/EcoSpold01" ...'. Such smaller sections will however not be valid anymore to import into for LCA software.

6.3.3 Merging separate XML files

Depending on the source, process inventories can be contained in their separate XML files, i.e. one process = one file. If a large amount of files shall be converted, it can be tiresome to use the standard procedure of having to select an individual XML file for the conversion of each single process. Here it would be advantageous to have a source XML file with *several* processes in them. This can be achieved in separate text editor applications like for instance TextWrangler, and many others. There commands exist that will convert either the contents of a whole folder or of selected files into one large text file. Look for commands like "Merge" or "Concatenate". In TextWrangler the procedure is straightforward: 1. Create a new empty document. 2. Select the command Edit → Insert → File Contents. 3) Select the files to be merged (Cmd button allows selection of non-adjacent files) 4. Save file with an ".xml" suffix and as DOS text file (cf. 6.3.1 on page 45). This new file can now be converted in Esixell in one go. Files created this way will parse well in Esixell, but it is important to know that they don't represent a valid ES1-XML file and probably won't import in LCA software⁴⁰.

You can still decide at this point, if you want to convert several processes into *one single* XL workbook, or have an *individual* workbook for each process, see setting preferences in chapter 3.3.1 'Split into several workbooks?' on page 12.

⁴⁰ For instance each individual source file contributes their header tag <ecoSpold> , which formally occurs only once per XML file. Or the sequential number given to datasets (ES1 field #200) will always be "1" instead of an increasing sequence of integers. These issues will naturally be remedied when back-converting XL→ML.

7 Conversion Excel to ES1-XML

The guide to convert an Excel file are described in chapter 'Excel to XML (XL to ML)' on page 8. With this conversion, the information found in the Excel cells are written into a newly created EcoSpold1-XML file (text file).

The Excel file needs to be in the proper layout (see chapter 5 on page 16). Esixell will test on presence of the four required sheets and it will check on presence of required cells, like for instance biosphere exchanges require a category and subcategory, or technosphere exchanges require a location.

7.1.1 Error Messages

If any of the *mandatory* requirements are not met,

- a) Esixell will display a dialog describing the error details,
- b) if possible highlight the erroneous cell in **orange**, select and display it in the workbook window
- c) and halt the creation of the XML file.

In case of errors, the XML file will not be valid or complete, and the mentioned error needs to be corrected first by the user, before the Esixell macro 'XLtoML' can be run again.⁴¹

Once errors in cells have been corrected, it is recommended that the **orange cell highlights** and cell comments inserted by Esixell are removed again for better visual guidance on remaining and upcoming errors. This can be done manually (since it is only one highlighted cell), but there is also a helper macro for that cf. chapter 4.1.1 'Macros to remove error formatting and comments' on page 14.

In case of workbooks with *multiple* datasets it is also recommended to consider whether a particular error is occurring also in other datasets, before attempting another conversion. Because with critical errors Esixell halts at the first encountered instance of an error and does not proceed any further.

7.1.2 Warning messages

In some particular cases Esixell will automatically provide default values when mandatory values are missing. This relates only to some cells of sheet 'X-Process'. The default values are listed in Tab. 5.3 on page 31. A warning message is generated, if such missing contents are encountered.

If the dates of the time period (startDate, endDate) are not given in the expected format yyyy-mm-dd, Esixell will try to expand information given only as years or year+months into full dates. If successful, this is also remarked in a warning comment.

A warning will also be created if a dataset has InfrastructureIncluded =1 but is apparently not a cumulated inventory (cf. chapter 5.10 on page 29).

A warning message will *not halt* the conversion (as an error would), but to notify the user a warning for a dataset is written as a *cell comment* into sheet 'X-Process' on the row proofReadingDetails (usually row 46). The cell *value* of the proofReadingDetails—i.e. the formal review comment—is not altered.

⁴¹ 'XL to ML' will halt conversion if important, mandatory information is missing, as to avoid creating any non-conform XML output files. In contrast the opposite operation 'ML to XL' will try to extract as much information as possible, even from a seriously defective XML file, so that required error correction can be performed centrally on one Excel file (and errors have not to be corrected in an unwieldy XML file in a text editor).

44		761	CountryCode		
45		762	PageNumbers	0	
46	Validation	5615	proofReadingDetails		* (database) version was missing in 'X-Process' for process acrylic render, at plantCH0kg ! (corrected to 0.00)
47		5616	proofReadingValidator		
48		5619	otherDetails		

Fig. 7.1 Non-critical warning text added by Esixell as cell *comment* (yellow box) on the line for review comments (proofReadingDetails). The cell *value* is unaltered, as to no interfere with any possible comments by the official, actual reviewer person.

7.1.3 Reviewer comments

As outlined in chapter 5.9 on page 28, the review information is not mandatory for Esixell. However, if a reviewer person is specified in sheet 'X-Process', also a non-empty review text is expected. Otherwise a parsing is interrupted and an error message is displayed.

7.1.4 Issues to check

Esixell cannot identify any and all problematic issues. One such issue is the occurrence of slightly deviating entries for actually identical entities.

Duplicate sources: Are the entries listed in "X-Source" sensible, or are there not required duplicates for identical reference sources? Check the columns in sheet "X- Source" for unnecessary duplicates of sources. See section 'Remaining problems' on page 35 for encountered examples of slightly deviating entries of actually identical sources. Esixell will neither be able to recognize those as identical, nor perform a correction.

To manually create a consolidated list (a) the duplicates must be deleted in "X-Source" and (b) the use of duplicate entries in activities must be consolidated as well. If for example the source with ID = 3 and ID = 8 are judged to be identical, a consolidated entry must be created in "X-Source" for instance of ID = 3, and then also all references to ID = 8 in sheet "X-Process" in row "ReferenceToPublishedSource" must be changed to ID = 3 (in that row a search and replace *for entire cell values* can be done to that effect). In "X-Process" the source with ID = 8 should be deleted for maximal clarity and avoidance of misunderstandings (but this is not strictly required; X-Source *can* contain sources *not used* in X-Process, even when they are duplicates, but keeping this is obviously inviting future inconsistencies in using the Excel file).

Duplicate persons: Are the entries listed in X-Person sensible, or are there not required duplicates for identical persons? Check the columns in sheet "X-Person" for unnecessary duplicates of persons. See section 'Remaining problems' on page 35 for encountered examples of deviating entries of actually identical persons. If a particular combination of person and organisation is listed more than once, incompatible references might result in different activities. Corrections to one entry might then erroneously not be reflected in activities referencing a duplicate entry. It is however accurate and desirable to list a particular person more than once when they have *changed* employers (cf. for example footnote 8 on page 24).

Consolidation of the listed persons works similarly as with sources: eliminating duplicates and making sure references to persons in "X-Process" are redirected to the remaining entries. Persons can be referenced in three locations in "X-Process": in row "DataEntryBy/Person" (ID 302), row "DataGeneratorAndPublication/Person" (ID 751), row "proofReadingValidator" (5616).

Allocated activities: Esixell can only process mono-functional activities. It will not parse multi-functional inventories. It is possible that certain datasets are derived from a multi-functional activity, like for instance rock salt electrolysis (NaCl) producing sodium hydroxide (NaOH) and chlorine gas (Cl₂) simultaneously. Changes in either of these three activities usually require a consistent adjustment in the other activities to maintain compatible inventories and a consistent allocation result. The user shall be aware of these dependencies and must make sure that any changes or corrections are applied consistently.

Inecoinvent 1.0-2.2 (2003-2010) the author would create an *unallocated* multi-functional activity where its meta text comments were copied into the derived, allocated mono-functional activities. The General comment should start with the text template "The multioutput-process '[MO-NAME]' delivers the co-products '[PRODUCT1]' and '[PRODUCT2]' and '[PRODUCT3]'". Therefore, activities related to each other by a performed allocation should ideally be identifiable via this uniform comment, but this cannot be guaranteed, especially for datasets not created and reviewed in the ecoinvent 2000 project.

8 Troubleshooting Esixell

8.1 Where are the converted files?

Esixell *tries* to save its exports into the same folder as the start file. Due to total path length limitations in MSExcel, this might not be possible (cf. chapters 8.3 and 8.2 below). Esixell will then try to *move up* the folder hierarchy automatically, until it can save the file with a short enough path. So look in the upper folders for recently created files. Consider moving the folder with your start files to a less deep location in the folder hierarchy.

8.2 Can't open XML File (ML→XL)

When converting an XML file into Excel, an error can occur in some Excel versions, saying that the file cannot be accessed ("runtime error '1004' Can't access file" or similar).

Convertible are only EcoSpold1 XML files and they must have the suffix ".xml". The '1004' issue can occur even for perfectly valid files, which have very long full file paths.⁴² The maximal file path length including the file name with its extension is 218 characters in some Excel versions. In XML files on deep subfolder levels or in folders with very long names, the maximal file path length limit can be exceeded, and Excel then refuses to open a file (even if the file itself is perfectly valid). Or even if an XL file can be accessed, a created Excel file might not be able to be saved in the target location (Esixell tries to save into the same folder as the start file), due to this path length limitation.

The solution here is to move the folder with start XML files to another location, with more shallow folder depth, or even the desktop.

8.3 Can't open Excel start file (XL→ML)

When converting an Excel file into XML, an error can occur in some Excel versions, saying that the start XL file cannot be opened. As in 8.2 above, the culprit is usually not an invalid file, but that the full file path is too long.

The solution here is to move the folder with start XL files to another location, with more shallow folder depth, or even the desktop.

8.4 Esixell toolbar still visible after closing Esixell

In some Excel versions deleting the toolbar upon close of Esixell is not possible. The toolbar is however assigned to be temporary only, which means the toolbar will be deleted, when Excel is closed.

The solution therefore is to quit Excel to delete the toolbar. Clicking on the close button of the toolbar will hide it, but not delete it.

Somewhat more intricately, the toolbar can also be removed manually. For this, choose the command "View/Toolbars/Edit..." or similar and there manually delete the toolbar. If you are experiencing this particular problem of lingering toolbars, it probably means that the *name* of the Esixell toolbar will be

⁴² The full file path includes the name of the drive, the entire folder hierarchy, and the file name and its suffix, and might look like this: "C:\Documents\Work\Project A\Folder 1\Subfolder 1.2\Subfolder 1.2.3\NewDatasetsES1.xml".

something like "user-defined 1" or similar (not "Esixell" as intended). Make sure to delete to correct toolbar. You can do this by *tooggling* the visible button in that dialog, before deleting.

9 Outlook

As this is the first public release of Esixell there are bound to be application problems and missing features. The author Gabor Doka welcomes any feedback regarding Esixell at do@doka.ch .

Possible future features might include a canonical check of technosphere and biosphere exchanges. I.e. checking if the given technosphere and biosphere exchanges are matching those defined in a separate MasterData list. This can avoid spelling and categorisation errors.

Another feature might be the creation of monofunctional ML process inventories from multifunctional XL inventories. This would require a XL template with additional columns (allocation factors) and additional programming. This can only be a *one-way conversion* "multifunctional → monofunctional". as the reverse is mathematically not possible, because the used allocation factors are not retrievable from the monofunctional inventories.

10 Appendix A: Meaning of EcoSpold1 integers

10.1 In sheet 'X-Exchange'

	Technosphere exchanges	Biosphere exchanges
InputGroup	5 = inputs of goods/services	4 = resources from nature
OutputGroup	0 = reference product 2 = by-product (only in multifunctional processes)	4 = emissions to nature
InfrastructureProcess	0 = false 1 = true	(not applicable)
UncertaintyType	0 = undefined 1 = lognormal (default) 2 = normal 3 = triangular 4 = uniform (rectangular)	(same codes as in technosphere)

10.2 In sheet 'X-Process'

ReferenceFunction	493	InfrastructureProcess	0 = false, 1 = true
DataSetInformation	201	Type	Always 1 = Unit process for UPR
DataSetInformation	202	Version	Ecoinvent database version, 0.00 for non-ecoinvent
DataSetInformation	203	energyValues	Not used, 0 = Undefined
DataEntryBy	304	QualityNetwork	Always 1 = ecoinvent, was never expanded
DataEntryBy	302	Person	ID number of person listed in sheet 'X-Person'
ReferenceFunction	400	DataSetRelatesToProduct	Always 1 = true
ReferenceFunction	494	InfrastructureIncluded	almost always 0 = false, see chapter 5.10 on page 29.
TimePeriod	603	DataValidForEntirePeriod	Almost always 1 = true
DataGeneratorAndPublication	751	Person	ID number of person listed in sheet 'X-Person'
DataGeneratorAndPublication	756	DataPublishedIn	0 = Data as such not published elsewhere, 1 = The data of some unit processes or subsystems are published, 2 = Data has been published entirely in
DataGeneratorAndPublication	757	ReferenceToPublishedSource	ID number of source/literature listed in sheet 'X-Source'
DataGeneratorAndPublication	758	Copyright	Usually 1 (= true)
DataGeneratorAndPublication	759	AccessRestrictedTo	Usually 0 = Public (open) For confidential UPRs: 1 = ETH Domain 2 = ecoinvent 2000 3 = Institute
Validation	5616	proofReadingValidator	ID number of person listed in sheet 'X-Person'

10.3 In sheet 'X-Source'

Sources	801	Number	ID number of source/literature to be used in field ReferenceToPublishedSource (#757) in sheet 'X-Process'
Sources	802	SourceType	0 = Undefined 1 = Article 2 = Chapters in anthology 3 = Separate publication (used often) 4 = Measurement on site 5 = Oral communication 6 = Personal written communication 7 = Questionnaires

10.4 In sheet 'X-Person'

Persons	5800	Number	ID number of person used in sheet 'X-Process' for author (Generator) secretary (Entry by) and reviewer (Validator).
---------	------	---------------	---

11 References

Web addresses indicate the source of electronic documents. The subsequent (date in brackets) refers to the date of retrieval.

- Ecoinvent 2002 Ecoinvent (2002) EcoSpold01 Schema Documentation - EcoSpold01Dataset_Simple.html. Swiss Centre for Life Cycle Inventories. Version of 29. Oktober 2002
http://www.ecoinvent.org/download/EcoSpoldSchema_v1.0.zip (29 Oct 2008) and identical at
https://www.ecoinvent.org/files/ecospold1_data_exchange_format.zip (1 March 2016, 27 Apr 2020)
- Hedemann & Meinshausen 2003 Hedemann, J, Meinshausen, I (2003) ecoinvent 2000 - Dokumentation EcoSpold. ifu Hamburg GmbH, 23. April 2003. <http://www.ecoinvent.ch/de/pdf/DokumentationEcoSpold1.2.pdf> (9 Mai 2003) EcoSpoldAccess Plugin available from <http://www.ecoinvent.org/database/older-versions/ecoinvent-version-2/methodology-of-ecoinvent-2/ecospold1/ecospold1.html> (11 Apr 2021)
- KBOB et al. 2018 KBOB (2018) KBOB-Ökobilanzdatenbestand 2016 und UVEK-Ökobilanzdatenbestand 2018. Koordinationskonferenz der Bau- und Liegenschaftsorgane der öffentlichen Bauherren KBOB, und eco-bau Nachhaltigkeit im öffentlichen Bau, and IPB Interessengemeinschaft Privater Bauherren, 7 Nov 2018. http://www.dflca.ch/inventories/Kommunikation_Oekobilanzdatenbestand_KBOB-2016_UVEK-2018.pdf (25 Nov 2018) see databases at https://db.ecoinvent.org/download/KBOB%20DQRv2_2016.zip?area=3e2c0806caa3c (25 Nov 2018) and https://db.ecoinvent.org/download/UVEK%20LCI%20Data_2018_20191003.zip?area=3e2c0806caa3c (2 Mai 2020)