Ref. Ares(2022)7429739 - 26/10/2022

ORIENTING

D3.2

Data specification

Specifications for a future LCSA data format

WP n° and title	**WP3 – LCSA Operationalisation**
Responsible Author(s)	**Thomas Sonderegger (ECA), Janire Clavell Diaz (TEC), Luca Zampori (PRE), Till M. Bachmann (EIF)**
Contributor(s)	**João S. V. Gonçalves (ECA), Simone Fazio (ECA), Marina Isasa (TEC), Mauro Cordella (TEC), Raul Ugarte (TEC), Mikel Astorquiza (TEC), Marco Bianchi (TEC), Rudri Mankad (PRE), Isadora Corrêa Hackenhaar (GHE)**
Dissemination level	**PU**
Version	**V2**
Keywords	**data, data specifications, data format, LCSA**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958231.

DOCUMENT INFORMATION

Grant Agreement Number	**958231**
Project Title	**ORIENTING: Operational Life Cycle Sustainability Assessment Methodology Supporting Decisions Towards a Circular Economy**
Status (F: final; D: draft; RD: revised draft):	**F**
Planned delivery date	**31/10/2022 (M24)**
Actual delivery date	**26/10/2022 (M24)**
Dissemination level: (PU = Public; PP = Restricted to other program participants; RE = Restricted to a group specified by the consortium; CO = Confidential, only for members of the consortium)	**PU**

DOCUMENT HISTORY

Version	Date (DD/MM/YYYY)	Description of changes	Contributors
01	**06/10/2022**		**ECA**
02	**26/10/2022**	**Updates considering contributions from partners**	**ECA, TEC, PRE, EIF**

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Tables

Figures

Acronyms

BoM	Bill of Materials
CF	Characterisation Factor
eILCD	extended International Life Cycle Data data format
EPLCA	European Platform on Life Cycle Assessment
GWP	Global Warming Potential
ILCD	International Life Cycle Data system
IPCC	Intergovernmental Panel on Climate Change
LC	Life Cycle
LCA	Life Cycle Assessment
LCC	Life Cycle Costing
LCI	Life Cycle Inventory
LCIA	Life Cycle Impact Assessment
LCSA	Life Cycle Sustainability Assessment
OIF	ORIENTING Input Format
OOF	ORIENTING Output Format
ORIONT	ORIENTING LCSA ontology
PSILCA	Product Social Impact Life Cycle Assessment database
S-LCA	Social Life Cycle Assessment
SHDB	Social Hotspot DataBase
XML	Extensible Markup Language

1. Executive summary

This report describes data specifications for a future LCSA data format following the methodological framework developed in the ORIENTING project (ORIENTING, 2022b). The aim is not to invent a new data format, but to discuss how to build on existing know-how and initiatives. Accordingly, data specifications are described in relation to the existing eILCD data format: Either existing data fields can be used to meet additional data needs or new fields would need to be added. Furthermore, the data specifications presented in this report are informed by and aligned with the ORIENTING LCSA ontology (ORIONT) presented in D3.1 (ORIENTING, 2022c). Converting these specifications into a complete new LCSA format would need a similar process as the development of the eILCD data format.

This report first discusses general data format requirements and introduces the eILCD data format. Then, extension needs for the eICLD format to arrive at an LCSA data format are discussed along the ORIENTING sustainability topics: (environmental) LCA, material criticality, material circularity, LCC, and social LCA.

Overall, it can be concluded that most additional data needs could one way or the other be satisfied with the existing data structure and data fields. Some additions are needed, in particular new flow datasets (having new/other flow types) and new flow property datasets (having new/other properties). For additional data to be stored and processed, the "Other" data field in eILCD datasets might be used for many things. However, the preferable solution would be to introduce new data fields with an appropriate name when adapting the eILCD data format to LCSA purposes.

A consensus data format (likely based on the eILCD format) and the software specifications provided in D3.3 (ORIENTING (2022d)), would expand the potential of integrating full LCSA in commercial software and can be seen as a first step to be considered in an ORIENTING exploitation strategy.

2. Introduction

This report describes data specifications for a future LCSA data format following the methodological framework developed in the ORIENTING project (ORIENTING, 2022b). The aim is not to invent a new data format, but to discuss how to build on existing know-how and initiatives. Accordingly, data specifications are described in relation to the existing eILCD data format (European Commission, 2022): Either existing data fields can be used to meet additional data needs or new fields would need to be added. A complete LCSA format would still need to be developed in a similar process as the eILCD data format was developed (see also section 2 in D3.3 (ORIENTING, 2022d)).

The data specifications presented in this report are informed by and aligned with the ORIENTING LCSA ontology (ORIONT) presented in D3.1 (ORIENTING, 2022c). ORIONT is a description of concepts used in LCSA and their relationships in general, considering ORIENTING needs (for example, the inclusion of material criticality and material circularity). Existing data formats, such as eILCD have an inherent structure, which is ontology-like. Hence, ORIONT has been elaborated considering the eILCD data format (see Figure 1).

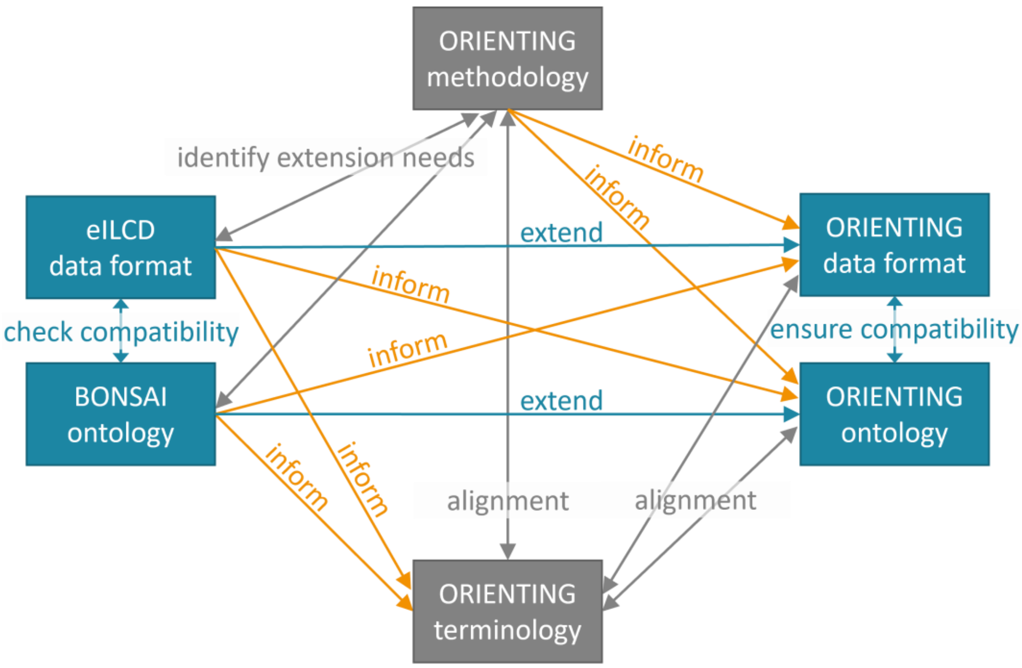


Figure 1. Material and workflow used to conceptualise the ORIENTING LCSA ontology (ORIONT) and data format.

This report is structured as follows: In section 3, a discussion on general requirements for the data format is presented. Subsequently, the eILCD data format is introduced in section 4, to afterwards discuss extension needs for each sustainability topic in ORIENTING: (environmental) LCA, material criticality, material circularity, LCC, and social LCA (S-LCA) in section 5. In section 6, the impact assessment is discussed. Finally, section 7 provides the concluding remarks to this deliverable.

3. General requirements for the data format

Generally, a data format needs to contain a) all information necessary for fulfilling the “tasks” intended to be performed, and b) meta data. In an LCA or LCSA, these tasks include:

- Storing information about processes/activities.
- Storing information about flows/exchanges including their properties.
- Modelling a (foreground) product system.
- Calculating and storing LCI results.
- Storing information about LCIA methods.
- Calculating and storing LCIA results.
- Displaying results.

Figure 2 displays the general information flow from data collection to result interpretation. For ORIENTING and its case studies, life cycle inventory information will be collected in (Excel) data collection sheets. These sheets might be further harmonized and become part of the project output. However, they (and their format) are not part of these data format specifications. The collected data and preparatory modelling output (for example, from an emission model) for which this data has been used, then needs to be brought into what we here call an “input format”, which can be read and used by software and tools to calculate results. The calculated outputs, on the other hand, need to be in an “output format”, which then can be read by the LCSA integration tool developed in Task 3.3 and documented in D3.4 (ORIENTING, 2022e).

Since the calculations of the results for the different topics covered in ORIENTING are likely to be performed with different tools, a common “output format” needs to ensure that these results are fed into the LCSA tool in a machine-readable way. Because this format is closely related to the software requirements specified in Task 3.2, the conceptualization of an ORIENTING Output Format (OOF) was developed within Task 3.2 (see section 2 in D3.3 (ORIENTING, 2022d)). Accordingly, this report mainly focuses on the ORIENTING Input Format (OIF). The two data formats might be integrated in the long run to specify a new future LCSA format (something like “eILCD+”) as indicated in Figure 2.

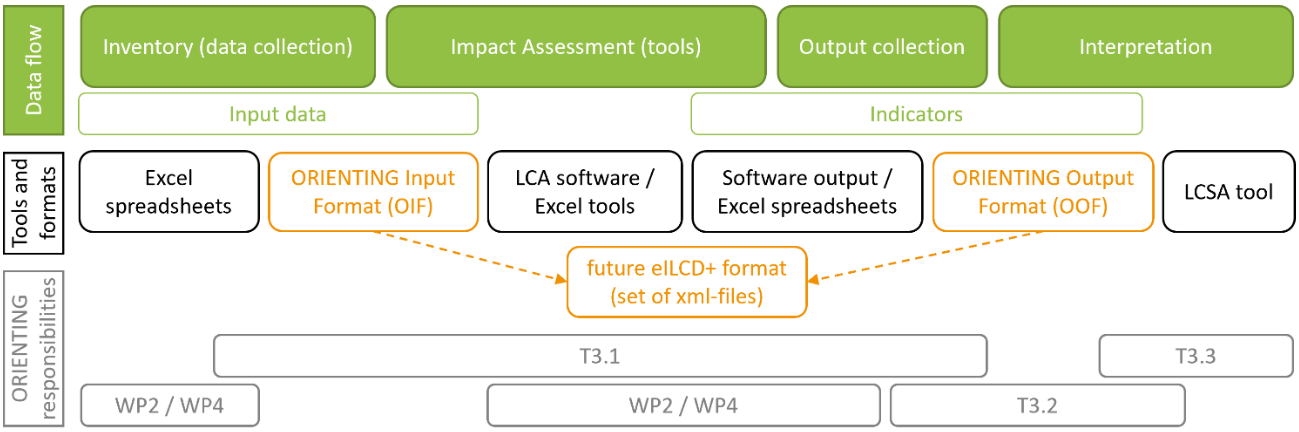


Figure 2. Data flow (from left to right), tools and formats, and ORIENTING responsibilities from data collection to result generation, integration, and interpretation; T: Task; WP: Work Package.

Because ORIENTING intends to draw on existing elements and notably on agreed state-of-the-art at EU level, the starting point for the OIF is the eILCD data format as this was developed by the European Commission and is linked to the Product Environmental Footprint (PEF). However, the eILCD data format was developed for the environmental domain only. Hence, it needs to be discussed how this format can be used to fulfil the needs of the ORIENTING LCSA methodology.

4. The eILCD data format

The eILCD data format is a collection of XML datasets. Compared to the ILCD data format, a further data set type, the life cycle model data set, was added to allow representing entire models consisting of ILCD process data sets (European Commission, 2022). The different datasets have an implicit hierarchical structure as shown in Figure 3.

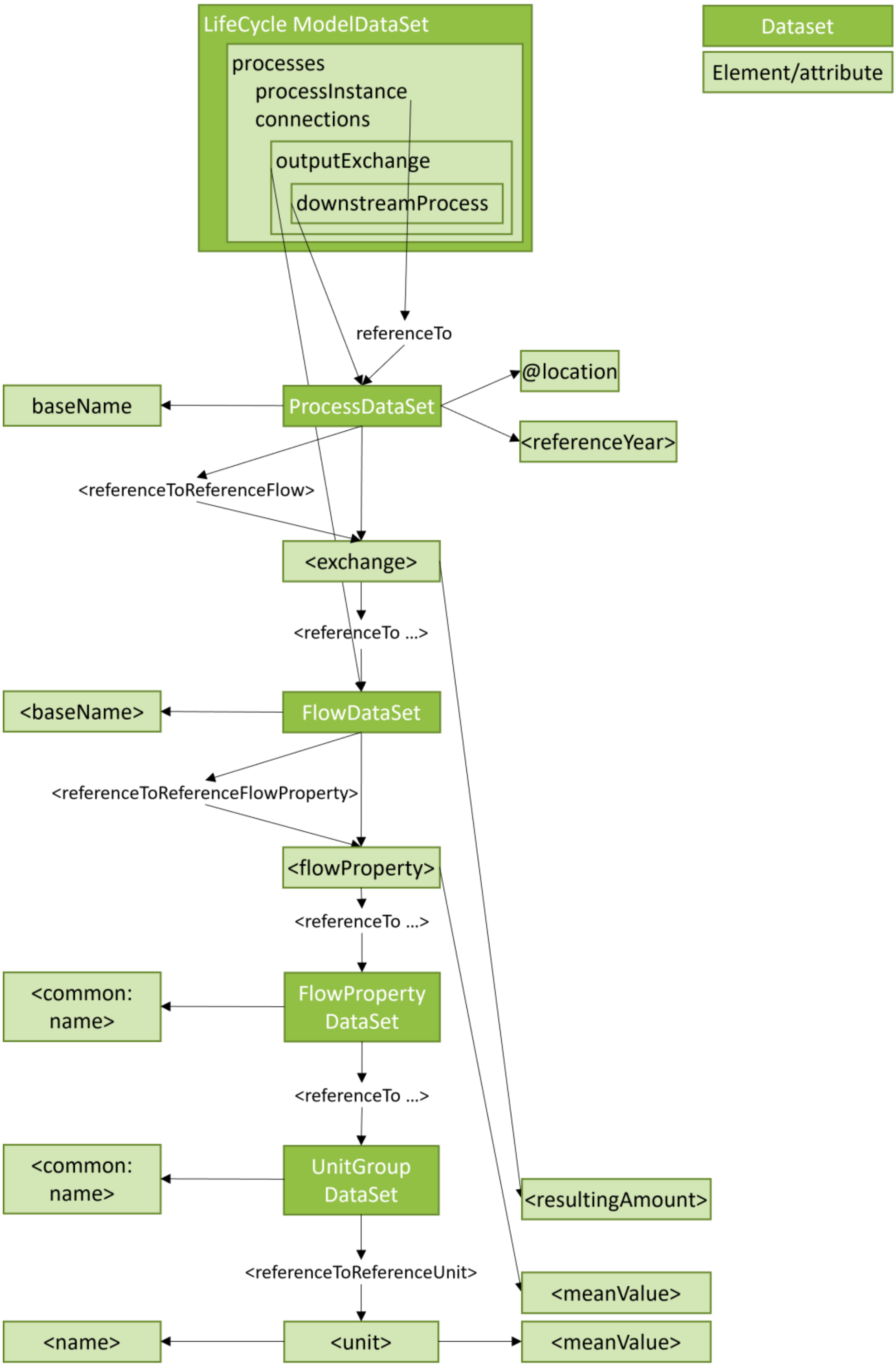


Figure 3. Basic hierarchical structure of the eILCD data format (European Commission, 2022).

As the name states, the Life Cycle Model dataset is a model that specifies processes included in a (foreground) product system as well as the connections between these processes and the background ones. Furthermore, the model dataset format allows to specify parameter settings of processes and assigning life cycle stages to processes. In a practical example: a Life Cycle Model dataset for “Packaging production” would specify how the process “Packaging production” is connected to several other processes through several flows (Figure 4). Furthermore, parameters in the transport dataset can be specified and life cycle stages could be assigned to the different processes, for example “Raw material acquisition” to “Plastic granulate production” and “Manufacturing” to “Packaging production”.

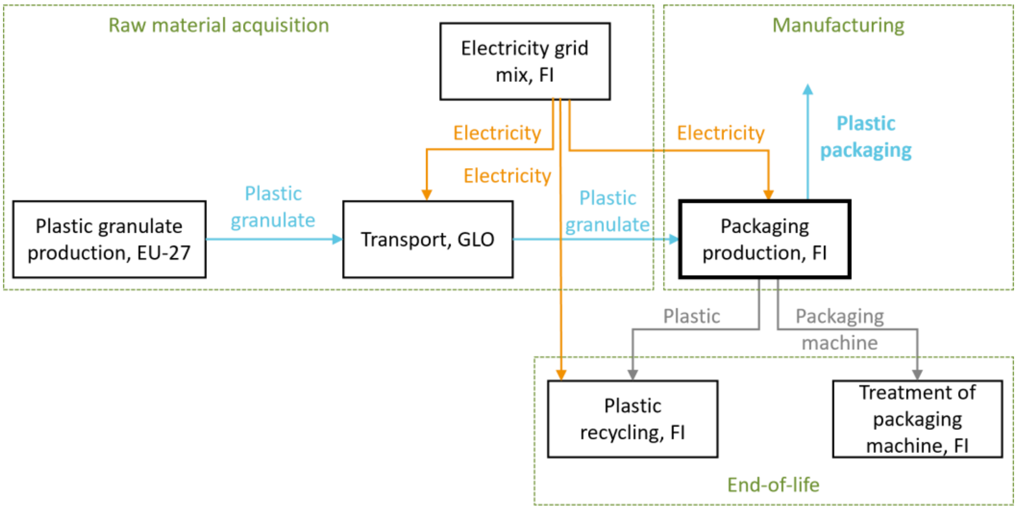


Figure 4. Example of an eILCD Life Cycle Model for “Packaging production, FI” (based on European Commission, 2022); FI: Finland; EU: European Union.

Processes are defined in dedicated datasets including the input and output flows (like in the older ILCD data format). These flows are then again defined in dedicated datasets including relevant flow properties. Flow property datasets then finally refer to unit group datasets, which contain the reference unit as well as other units that might be used for conversion (Figure 3 and Table 1). The connection between the datasets follows the hierarchical structure and is established with “referenceTo…” XML tags in the files (see Box 1). For example, an input or output flow in a process dataset contains a reference to a flow dataset. Thereby, the reference flow, property, and unit which the dataset refers to is defined by “referenceToReferenceFlow”, “referenceToReferenceFlowProperty” and “referenceToReferenceUnit” tags in the single datasets.

Box 1. XML terminology.

Tag: A tag begins with “<” and ends with “>”

Element: An element either begins with a start-tag and ends with a matching end-tag such as `<greeting>Hello, world!</greeting>`, or it consists only of an empty-element tag such as `<line-break/>`.

Attribute: An attribute is a name–value pair within a start-tag or empty-element tag such as `<img src="cat.jpg" alt="Cat" />`, where the names of the attributes are "src" and "alt", and their values are "cat.jpg" and "Cat" respectively.

Source: https://en.wikipedia.org/wiki/XML#Key_terminology

In the eILCD data format documentation (https://eplca.jrc.ec.europa.eu/LCDN/developerILCDDataFormat.xhtml), attributes start with “@”. It is possible to include information as an element or as an attribute. The value of a flow, for example, could be included as an element (`<value>1</value>`) or as an attribute (`<flow value=”1”…>`)

Some important datasets constituting the eILCD data format are shown in Table 1, Table 2, and Table 3. The full format can be downloaded from the EPLCA website (European Commission, 2022). Files listed in Table 1 describe dataset structures, meaning that collected data for each process, for example, needs to be brought into a file following the data structure given in the "ILCD ProcessDataSet" file. The number of these files depends on the number of processes, flows, etc. in the LCA model.

Table 1. Datasets of the eILCD data format.

File	Description / "eILCD definition"
ILCD LifeCycleModelDataSet (eILCD)	Contains the modelled system, meaning the (references to) processes and the flows connecting them. "System of interconnected processes that jointly represent the full or partial life cycle of a product system, with the final process step ('Reference process instance') being referenced from this life cycle model data set, which connects and stepwise scales all other process instances via the connecting product and waste exchanges."
ILCD ProcessDataSet	Contains process data including the (references to) input and output flows "Data set for unit processes, partly terminated systems, and LCI results. May contain LCIA results as well."
ILCD FlowDataSet	Contains flow data including the (references to) flow properties.
ILCD FlowPropertyDataSet	Contains the flow property data including (references to) unit groups.
ILCD UnitGroupDataSet	Contains unit information including conversion factors from the reference unit of the dataset to other units.
ILCD LCIAMethodDataSet	Contains the LCIA method data including the CFs with references to the flow datasets.
ILCD SourceDataSet	"Data set for bibliographical references to sources used, but also for reference to data set formats, databases, conformity systems etc."
ILCD ContactDataSet	Contains data about people and institutions involved in the data generation

Table 2 describes important "stylesheets". These are datasets containing information specified by the data format such as the categorization system for flows or the list of available locations in the ILCD data system. These datasets exist only once.

Table 2. Important "stylesheets" of the eILCD data format.

File	Description
ILCDClassification_Reference	Contains classification systems organized in levels for different elements in the ILCD data system, for example, processes, flows, or LCIA methods.
ILCDFlowCategorization_Reference	Contains the categorization organized in levels for elementary flows, for example, "Emissions", "Emissions to water", "Emissions to fresh water".
ILCDLCIAMethodologies_Reference	Contains names of LCIA methods.
ILCDLocations_Reference	List of locations available in the ILCD data system.

Table 3 describes important "schemas". These are datasets containing information about data types and enumeration values used in the ILCD data system. These datasets exist once.

Table 3. Important "schemas" of the eILCD data format.

File	Description
ILCD_Common_DataTypes.xsd	Definition of data types used in the ILCD data system such as "Year" defined as "xs:integer" and "4-digit year".
ILCD_Common_EnumerationValues.xsd	Prescribed lists of values for certain data fields such as "LCIAImpactCategoryValues" containing the possible LCIA impact category names.

Considering the tasks to be performed, mentioned in the beginning, Table 4 shows in which eILCD datasets this can be done. Accordingly, these will also be the main datasets that might need some further extensions for being ORIENTING LCSA compatible (see next section).

Table 4. Mapping data format tasks to eILCD datasets

Task	Dataset
Storing information about processes/activities	ILCD ProcessDataSet
Storing information about flows/exchanges including their properties	ILCD FlowDataSet, ILCD FlowPropertyDataSet
Modelling a (foreground) product system	ILCD LifeCycleModelDataSet (eILCD)
Calculating and storing LCI results	storing: ILCD ProcessDataSet
Storing information about LCIA methods	ILCD LCIAMethodDataSet
Calculating and storing LCIA results	storing: ILCD ProcessDataSet
Displaying results	based on ILCD ProcessDataSet or ILCD LifeCycleModelDataSet (eILCD)

5. Extension needs for an ORIENTING LCSA data format

The eILCD data format was developed in the context of the PEF framework. Therefore, its main goal was to support environmental LCA, while other sustainability topics were out of scope. To make the eILCD data format fit to a broader LCSA scope, extensions are needed to capture LCSA as defined in ORIENTING, including LCC and S-LCA as well as material criticality and material circularity assessments. The European Commission is coordinating the efforts around further developing the ILCD and the eILCD data format. To make such extensions official in an updated eILCD data format would, therefore, require the commitment from the EC, time, and consensus building. This means that this is only a mid-term to long-term solution. Within ORIENTING, a practical solution, applicable in the short-term, was needed. Work in ORIENTING therefore includes:

a) the discussion of what an ORIENTING Input Format (OIF, see section 3 and Figure 2) could look like starting from the existing eILCD data format for environmental LCA (this report and in particular the following sections); and

b) the development of an ORIENTING Output Format (OOF, see D3.3 (ORIENTING, 2022d) to enable an easy result compilation by companies and a well-defined structure to enable a correct import into a software environment (such as the ORIENTING LCSA tool developed in Task 3.3 and documented in D3.4 (ORIENTING, 2022e)).

The discussion of point a) along the ORIENTING sustainability topics is provided in the following sections. Furthermore, overarching topics are discussed in more detail.

5.1. LCA

The eILCD data format was conceptualized for LCA and can be considered generally fit for purpose. Two methodological demands affect LCA: the assessment of life cycle stages (see section 5.6) and regionalisation (see section 5.7).

5.2. Material criticality

Material criticality refers to raw materials. In ORIENTING, these include metals (as single elements), minerals (such as magnesite) and aggregates (such as silica sand), as well as fossil fuels and natural biomass. In LCA terms, these can be elementary flows or intermediate flows. Most of the EU raw material list (Blengini et al., 2020) can be mapped to elementary flows in the ecoinvent database (v3.8; Wernet et al., 2016) (Table 5). However, some materials only exist as intermediate flows: hydrogen (gaseous or liquid), phosphorus (white, liquid), coke, limestone (unprocessed), and silica sand. The amounts of intermediate flows are not part of the standard LCI result, which only displays elementary flows. However, they could theoretically be retrieved from LCA calculations. This would need some adjustments on the software side. When it comes to storing results, one possible solution would be to transform these flows to “artificial” elementary flows. In that case, this would not need an adjustment of the data format as they could be stored in the “exchanges” element of a process dataset.

Table 5. Mapping of the EU raw material list (Blengini et al., 2020) to elementary flows in the ecoinvent database (v3.8; Wernet et al., 2016); colour code explained at the bottom of the table

Raw Material	Comment	Name
Aggregates	sand	Sand, unspecified, in ground
Aggregates	gravel; gravel round or crushed as intermediate flows	Gravel, in ground
Aluminium		Aluminium, in ground
Antimony		Antimony, in ground
Arsenic		Arsenic, in ground
Barytes		Barium, in ground
Bauxite	bauxite is analysed in the EU report as the main source of aluminium. The factsheets are presented jointly	Aluminium, in ground
Bentonite		Clay, bentonite, in ground
Beryllium		Beryllium, in ground
Boron/Borate	(sodium borate) "refined borax" as one of the processed products that contains boron	Borax, in ground
Cadmium		Cadmium, in ground
Cerium		Cerium, in ground
Chromium		Chromium, in ground
Cobalt		Cobalt, in ground
Coking Coal	hard coal has a carbon content of about 92-98%	Coal, hard, unspecified, in ground
Copper		Copper, in ground
Diatomite		Diatomite, in ground
Dysprosium		Dysprosium, in ground
Europium		Europium, in ground
Feldspar		Feldspar, in ground
Fluorspar		Fluorspar, in ground
Gadolinium		Gadolinium, in ground
Gallium		Gallium, in ground
Gold		Gold, in ground
Gypsum		Gypsum, in ground
Hafnium		Hafnium, in ground
Hydrogen	According to the EU factsheet report 2020: "Hydrogen is not found free in nature and must be “extracted” from diverse sources: fossil energy, renewable energy, nuclear energy and the electrolysis of water." i.e., i.e., intrinsically a intermediate flow.	NOT CHARACTERIZED
Iron ore		Iron, in ground
Kaolin clay		Kaolinite, in ground
Lanthanum		Lanthanum, in ground
Lead		Lead, in ground
Limestone	calcite is composed of calcium carbonate	Calcite, in ground
Lithium		Lithium, in ground
Magnesite		Magnesite, in ground
Magnesium		Magnesium, in ground
Manganese		Manganese, in ground
Molybdenum		Molybdenum, in ground
Natural cork		Wood, unspecified, standing
Natural graphite		Metamorphous rock, graphite containing, in ground
Natural rubber	i.e. latex or 2-methylbuta-1,3-diene; (C5H8)n;	NOT CHARACTERIZED
Natural teak wood	Teak is a hardwood tree species	Wood, hard, standing
Neodymium		Neodymium, in ground
Nickel		Nickel, in ground
Niobium		Niobium, in ground

Raw Material	Comment	Name
Palladium		Palladium, in ground
Perlite		Perlite, in ground
Raw Material	**Comment**	**Name**
Phosphate rock		Phosphorus, in ground
Phosphorus	According to the EU factsheet report 2020: "Elemental phosphorus here refers to the specific forms of the element phosphorus (P) in which it is produced as an isolated element (P4) in dedicated electrothermal reducing furnaces (in different forms: white/yellow or red phosphorus)." i.e., intrinsically a intermediate flow	
Platinum		Platinum, in ground
Potash		Potassium, in ground
Praseodymium		Praseodymium, in ground
Rhenium		Rhenium, in ground
Rhodium		Rhodium, in ground
Samarium		Samarium, in ground
Sapele wood	Sapele is a hardwood tree species	Wood, hard, standing
Scandium		Scandium, in ground
Selenium		Selenium, in ground
Silica sand	diatomite is cited as an source of silica sand in the EU report	Diatomite, in ground
Silica sand	Perlite is cited as an source of silica sand in the EU report	Perlite, in ground
Silicon metal		Silicon, in ground
Silver		Silver, in ground
Strontium		Strontium, in ground
Sulphur		Sulfur, in ground
Talc		Talc, in ground
Tantalum		Tantalum, in ground
Tellurium		Tellurium, in ground
Terbium		Terbium, in ground
Tin		Tin, in ground
Titanium		Titanium, in ground
Tungsten		Tungsten, in ground
Vanadium		Vanadium, in ground
Yttrium		Yttrium, in ground
Zinc		Zinc, in ground
Zirconium		Zirconium, in ground
LREEs		
HREEs		
PGMs		
missing match		
questionable match		
Upcoming in v3.9:		
Bismuth		Bismuth, in ground
Germanium		Germanium, in ground

Even if accessing intermediate flows in LCI results could become possible, two methodological questions remain: a) to which amount of metal/raw material the characterization factors should be applied, and b) whether the current data structure works or whether changes in data format would be required. Different options to assess material criticality are provided in Table 6 and their application potential is assessed below.

Table 6. Options to assess material criticality depending on amounts and material considered; BoM: Bill of Materials, LCI: Life Cycle Inventory.

Amounts considered	**Materials considered**	
	Only materials in BoM	*All materials in LCI*
Amounts in product	Option 1	
Amounts in product + losses in supply chain	Option 2	
Amounts in product + losses in supply chain + indirect use	Option 3A	Option 3B
Amounts in LCI	Option 4A	Option 4B

One option to assess material criticality is to simply use the bill of materials (BoM) and hence the amounts (of elements, minerals, and aggregates) in a product (option 1 in Table 6). Another option is to use the full LCI result of a product (option 4B). However, several aspects need to be considered for the latter option: First, this does not only include metals in the product, but also those used somewhere in the background, for example in electricity production. Second, the amounts might be lower than the amounts in the product because economic allocation is systematically applied to multi-output processes in the ecoinvent cut-off model (see Box 2 for an explanation of the issue). All other options in Table 6 cannot be simply taken from the BoM/product composition or the LCI result and hence would need additional modelling.

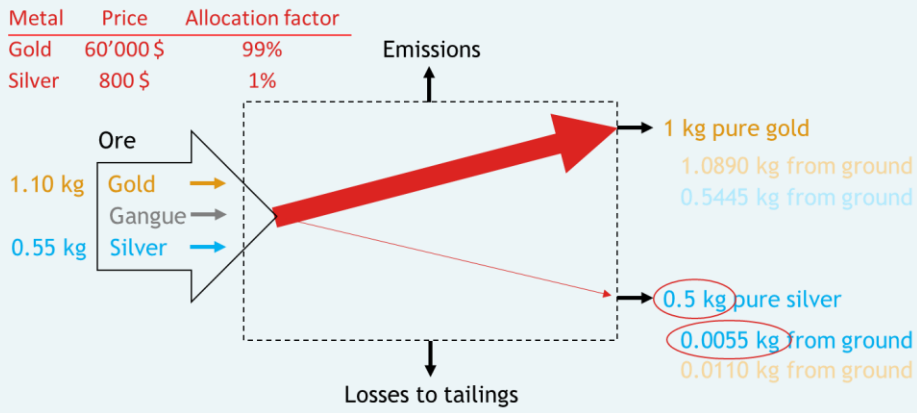
Options 3A and 3B would be possible with an adjusted allocation mechanism for elements/metals that would maintain physical flows to products and only economically allocate the losses (Sonderegger, 2021). Option 2 would require additional calculations along the supply chains, making sure that only losses and not indirect uses are considered, which could become a quite complicated procedure. Adjusted allocation and tracking supply chains both might be generally achieved if metal contents in products can be tracked throughout the data, however, this is something to be explored and not yet available as a result from databases and software.

The two most feasible options for now are 1 (amounts in product) and 4B (full LCI). From the data format point of view, no changes are needed for the data structure to implement material criticality according to these options, including the impact assessment. For option 1, the list of elemental contents in the flow property datasets needs to cover the list of critical raw materials, which is not yet the case.[1] Furthermore, CFs need to be applied to these elemental contents belonging to a flow, so to a flow dataset. The format does not need to be changed for that. Option 4B is easy because the impact assessment can be applied to the same LCI result as for the LCA. However, it comes with some distortions as explained above and in Box 2.

1 https://eplca.jrc.ec.europa.eu/permalink/EF3_1/EF-LCIAMethod_CF(EF-v3.1).xlsx; accessed 2022-10-04. The sheet "FlowProperties" contains flow properties including metal contents, a comparison to Table 5 shows that dysprosium and hafnium content are missing.

Box 2. Effects of economic allocation on amounts in inventory results.

With the application of economic allocation, environmental flows including metal flows from ground contained in ores are allocated based on the value of the metals. This can result in lower flows from ground than contained in the product. This is not wrong, but only reflects economic reality (what drives the activity by producing revenue) and not physical flows. In case of material criticality, this might not be the flow to be assessed, but this is a methodological choice.



LEARNINGS FROM MATERIAL CRITICALITY FOR DATA FORMAT

- The main feature in the data format that allows for tracking elements is the “flow properties”. These need to be complete regarding elemental contents, meaning they should at least cover elements covered in the critical raw materials reports. Ideally, they would cover the full periodic table.
- The currently available ILCD flow properties need to be extended with missing elements.
- To make all flows (elementary or intermediate) showing amounts in the product plus losses in the supply chain, a customized solution would be necessary as LCI results do not show that. However, if elemental contents of flows are known, this should be possible, so no need to adjust the data format.
- Showing other than elementary flows in LCI results might need a transformation of these flows into “artificial” elementary flows. Then these would fit into the data format.

5.3. Material circularity

Data needs for circularity indicators are product related and include, for example, mass, lifetime, recycled/secondary input/feedstock, recyclable output, and recycling efficiencies (Figure 5, see section 12.2 in D2.3 (ORIENTING, 2022b) for details). Thereby, data is mostly retrieved from companies, some average data on the product analysed is needed as well (D2.3). Hence there is a mix of process “properties” and flow properties, and circularity data could be integrated at different levels: at the process level (ILCD ProcessDataset) and at the product level (ILCD FlowDataset). ecoinvent has faced a similar integration challenge when introducing the EN15804 system model (Ioannidou et al., 2021), which includes similar indicators (for example, “Materials for Recycling”) and data needs that are not based on elementary exchanges. Thereby, most of the indicators are modelled as “artificial” elementary exchanges or properties, depending on whether the indicator is a characteristic of the process or the product (Ioannidou et al., 2021). The same approach could be followed for storing circularity data in the eILCD data structure. Mass and lifetime, for example might be considered product/flow properties, for which a new flow property dataset could be introduced, whereas recyclable input and output could be modelled as additional data point in the process dataset. The section that seems most suitable is the exchanges section where data points could be modelled as inputs and outputs. They do not need to be “artificial” elementary exchanges, but could also be categorised as “other”, which is the category intended for “modelling support flows”[2].

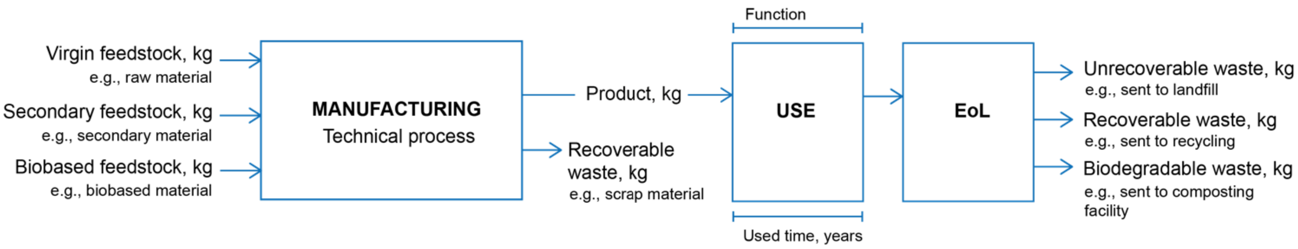


Figure 5. Example of data collection needs for a material circularity assessment (ORIENTING, 2022b).

With a mix of new flows and flow properties, a software would need to be able to process this data points jointly to calculate circularity indicators. If one would like to go closer to LCA, where only flows are assessed, the needed flow properties would need to be modelled as “other” flows as well. Which approach would be preferable is something that would need to be discussed with software providers. Another thing to be further examined is how to deal with multi-output processes and whether some of the data (such as recycled input) would need allocation to different products.

LEARNINGS FROM MATERIAL CIRCULARITY FOR DATA FORMAT

- Data points needed for calculation of circularity indicators could be introduced as exchanges (“other” flows) in process datasets (which requires corresponding flow datasets) and as flow properties in new flow property datasets.
- The currently available ILCD flows and flow properties need to be extended.

[2] https://eplca.jrc.ec.europa.eu/LCDN/downloads/ILCD_Format_1.1_Documentation/ILCD_Common_EnumerationValues.html#FlowTypeValues; accessed 2022-10-05

5.4. LCC

The collection of economic data is different for foreground and background activities. In the case studies, many different cost elements for an activity and resulting products can be collected (preferably based on primary data, cf. Cost Breakdown Structure (CBS) presented in D2.3 (ORIENTING, 2022b)). However, this is not applicable to a full background database. There, the prices of products need to be collected from other sources (largely secondary data).

For the foreground system, the cost elements accounted for in the calculation might be much more detailed and comprehensive and include costs related to activities (such as labour, maintenance, management, or training) and non-tangible goods (such as licenses, trademarks, or renting) (Sonderegger et al., in preparation). Which costs/prices should be included in the end in inventory datasets depends on what indicators are intended to be calculated. In any case, these different elements can be attached to products as flow properties. For this, the currently available ILCD flow properties would need to be extended as they do not yet include any economic information. Guidance on what data to collect and how to collect it to have reliable cost/price/revenue data for the foreground and the background is provided in D2.3 (ORIENTING, 2022b).

LEARNINGS FROM LCC FOR DATA FORMAT

- Cost/price/revenue data per product can be stored as a flow property.
- The currently available ILCD flow properties would need to be extended to include cost/price/revenue information.

5.5. Social LCA

Generally, the granularity of assessment differs a lot for foreground and background. As described for social LCA (S-LCA) in D2.3 (ORIENTING, 2022b), primary data might be collected for “those processes where the company has some control over, or from processes or companies, with which the company has established relationships”. This means that, for example, several suppliers within one or different countries can supply the same intermediate product (see Figure 6) and all these suppliers would get a supplier-specific assessment (“supplier” meaning the entity performing a process/activity, for example, a specific factory in a certain country ran by a company). However, this option is not yet explicitly part of existing data formats, but an activity is usually defined by its name and its location (usually a country, sometimes a region). For the background or the remote parts of the life cycle, secondary data is needed. The “traditional” S-LCA databases (SHDB and PSILCA, see D1.2 (ORIENTING, 2022a)) that can be used for that do not have a supplier-specific but a country/region-sector resolution.

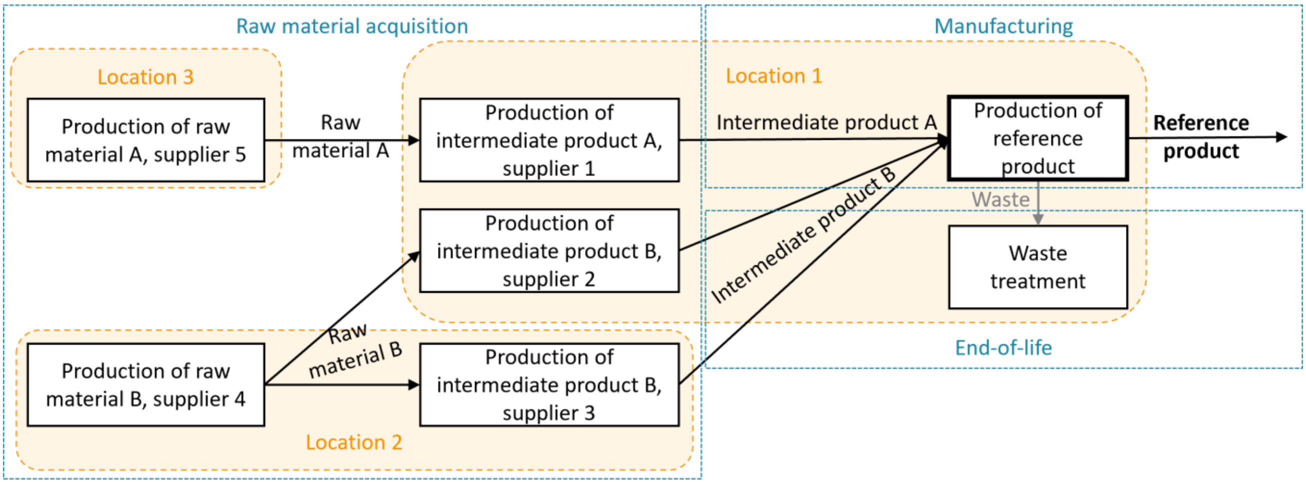


Figure 6: Conceptualization of a product system with different suppliers in different locations sometimes delivering the same products.

The “supplier issue” relates to the topic of “regionalisation” (section 5.7). The “geography” element in the process dataset of the eILCD data format already has an “location” and an “latitudeAndLongitude” attribute. Locations are countries (sometimes states) or regions. While this resolution might be too coarse to distinguish suppliers, providing coordinates might be more precisions than needed (although coordinates could be attributed to production facilities). Another option would be to use the “other” element in the “geography” element could be used to provide an identifier for a specific supplier. This information might also be used in the name to help identifying a process by name.

The method for S-LCA suggested for ORIENTING is a qualitative and semi-quantitative reference scale approach, meaning that social performance is assessed on a scale from -2 to 2 (negative impact to positive impact) (see section 21 in D2.3 (ORIENTING, 2022b)). The inventory data collection might therefore be considered also the collection of impact assessment data as it entails “collecting information, evidence and supporting information (or data) on the social topics identified in the goal and scope of the study” (D2.3 (ORIENTING, 2022b)). This information is structured along impact categories/social topics which affect certain stakeholders (see Figure 11 in D3.3 (ORIENTING, 2022d)). Social performance scores per social topic and stakeholder could be included as inventory information in the “exchanges” element of process datasets. Thereby, “other” flows per social topic and stakeholder would need to be created to allow for unique identification. The values of these flows would be independent of the reference flow, meaning that the values would not be scaled with the amount of reference flow.

An important information needed to connect to background databases, which have a country/region-sector resolution is the classification of processes and products. PSILCA is classifying activities with the UN ISIC classification (UNstats, 2022a, Maister et al., 2020). For use of the SHDB, “materials used to produce a product or purchases made by an organization” need to be attributed to one of the 57 GTAP sectors (Benoit-Norris et al., 2019). ecoinvent, as an example of an LCA database, uses ISIC for processes/activities and UN classification CPC for products (UNstats, 2022b). In eILCD, classifying processes is possible in the “Classification” element of the “Process data set” and classifying flows is possible in the “Classification” element of the “Flow data set”. In the end, the classification of activity and product (flow) should be aligned. Theoretically, there is the exception where a by-product belongs to a different sector than the activity and its main product.

LEARNINGS FROM S-LCA FOR DATA FORMAT

- S-LCA in ORIENTING needs supplier-specific information, which can go beyond the resolution given by activity and location (country or region). Several solutions are possible within the existing data format. However, the introduction of a new data field is of course also possible.
- Categorization of activities and products according to economic/industry sectors is possible with the existing format.
- Inventory information is provided in a qualitative and semi-quantitative way on a scale from -2 to 2. This information might be included in the "exchanges" element of process datasets.

5.6. Assessing life cycle stages

To show hot spots along the life cycle (LC), results need to be available per LC stage. This means that these stages need to be assigned to the different processes of the analysed system. In the eILCD data format, LC stages are specified in the life cycle model dataset ("groupDeclarations" element). The European Commission (2021) suggests the following LC stages as the minimum in a PEF study:

1) raw material acquisition and pre-processing (including production of parts and components);
2) manufacturing (production of the main product);
3) distribution (product distribution and storage);
4) use;
5) end of life (including product recovery or recycling).

For ORIENTING, two stages ("Design - R&D" and "Maintenance, repair, refurbishment") were added as they might be important for some topics (notably material circularity). The naming was slightly adjusted (see, for example, Figure 16 in D2.3 (ORIENTING, 2022b)):

1) Design - R&D
2) Raw material acquisition
3) Manufacturing
4) Installation/distribution/retail
5) Use
6) Maintenance, repair, refurbishment
7) End-of-life

If LC stages are declared in the life cycle model dataset, organising results by LC stages depends on whether the software used is able to do this.

LEARNINGS FROM LIFE CYCLE STAGES FOR DATA FORMAT

- Life cycle stages can be declared in the life cycle model dataset.

5.7. Regionalisation

The main methodological extension of the environmental assessment within ORIENTING is land use related. Three indicators will be added: 1) an updated soil quality index based on three LANCA indicators (erosion resistance [kg soil], mechanical filtration [m^3 water] and groundwater replenishment [m^3 groundwater]), 2) biotic production [kg biotic production], and 3) a biodiversity indicator (ORIENTING, 2022b). Furthermore, an improved framework for land use impact assessment is developed that considers three levels of detail to model and assess land use activities (ORIENTING, 2022b):

1. Country- (or region-) average land use flows assessed with country- (or region-) average CFs;
2. Geography-specific land use flows assessed with CFs from maps (coordinates or GIS vector file);
3. Calculation of foreground-specific CFs based on primary data that includes details such as management practices and soil quality parameters that are used as input in LANCA.

The geography-element in the process dataset already has a "latitudeAndLongitude" attribute, which can contain coordinates. There is no data field for a reference to a GIS vector file. However, there is the "other" element, which could be used for this. Furthermore, there is a subLocationOfOperationSupplyOrProduction-element, which could be used for specifying "sampling sites of a company-average data set, the countries of a region-average data set, or specific sites in a country-average data set"[3]. If flows inherit the geography from a process in the software being used, then this information needs to be stored in the calculated LCI results so the flows with the different geographies can be assessed with the corresponding CFs (Figure 7). The eILCD process dataset only has a "location" element for flows, but not "latitudeAndLongitude" attribute within that. This would need to be aligned or again the "other" element which could be used.

Assuming that in the example in Figure 4 "Plastic granulate production" and "Packaging production" both need water, the LCI for "Packaging production, FI" should contain water flows for (at least) EU-27 and Finland, so the LCI result would look like shown in Table 7.

Table 7. Example for a life cycle inventory result at country/region resolution.

Plastic packaging production, FI			
Product flow	Plastic packaging	1	kg
Elementary flow	Water, EU-27	3	m^3
Elementary flow	Water, FI	10	m^3

Assuming further that the site in Finland is known by coordinates, the LCI result would look like shown in Table 8.

Table 8. Example for a life cycle inventory result at country/region and coordinates resolution.

Plastic packaging production, FI			
Product flow	Plastic packaging	1	kg
Elementary flow	Water, EU-27	3	m^3
Elementary flow	Water, (61.50, 23.75)	10	m^3

[3] https://eplca.jrc.ec.europa.eu/LCDN/downloads/ILCD_Format_1.1_Documentation/ILCD_ProcessDataSet.html; accessed 2022-10-06

The list of elementary flows does not need to be extended as it does not contain geography information. However, the data format needs to be able to contain and process the location information for the inventory result and the CFs (Figure 7).

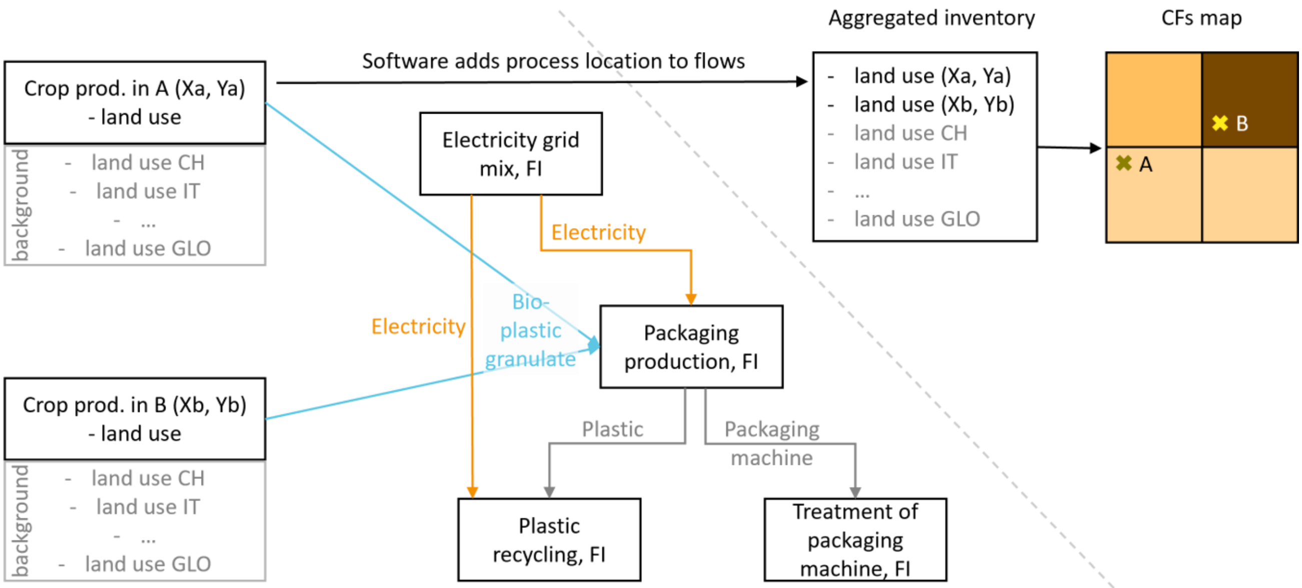


Figure 7. Conceptual visualization of how location information in coordinates would be transferred to LCI results.

LEARNINGS FROM REGIONALISATION FOR DATA FORMAT

- The eILCD process data set contains a data field for coordinates.
- There is no data field for a reference to a GIS vector file. The “Other” field could be used.
- For the exchanges in the process dataset, there is no such element or attribute. The “Other” field could be used.

6. Impact assessment and the ORIENTING Output Format (OOF)

The Orienting Output Format (OOF) generally is organised by sustainability topic and resolved by indicators and LC stages (and the total over all LC stages).[4] This is also how the results would need to be made available by a software and the (future) data format. The "LCIA results" section in the eICLD process dataset only has the "LCIA method" field (referring to a LCIA method dataset) to provide information about what impact is quantified. This means that an LCIA method dataset cannot contain a method with different indicators (IPCC 2021 with GWP20 and GWP100), but only one specific indicator of a method (IPCC 2021 GWP 100). This is confirmed by the definition of the "Name of LCIA method or methodology" data field in the dataset:

> Name of the data set. Composed as follows "LCIA methodology short name; Impact category/ies; midpoint/endpoint; Impact indicator; Source short name". Not applicable components are left out. Examples: "Impacts2007+; Climate change; midpoint; Global Warming Potential; IPCC 2001"; "ABC 2006; Acidification; endpoint; Species diversity loss; John Doe 2006"; "My-indicator2009; combined; endpoint; Ecopoints; various"[5]

The example given is not clear (only stating "Global Warming Potential"), but the naming structure clearly states "impact indicator". Hence, for each indicator in the OOF an LCIA method dataset needs to exist. These datasets might not contain CFs as not all LCSA impact indicators are calculated in the way it is done for LCA, meaning by multiplying elementary exchanges with CFs. However, the datasets should still include all information about the method and indicator used. With this, results per indicator can be stored in a process dataset. For S-LCA this would mean to produce one LCIA method dataset per social topic-stakeholder combination (see section 5.5) as for now the aggregation or presentation of these indicators are not clear (see Table 7 in D3.3 (ORIENTING, 2022d)). To group LCIA scores per sustainability topic, methods need to be classified in some way. One option would be to use the "classification" element in the LCIA method dataset. Furthermore, this could also be used to group indicators within S-LCA by social topic or stakeholder group. Organising results per LC stage can be done as described in section 5.6.

LEARNINGS FROM THE OOF FOR DATA FORMAT

- For each indicator, an LCIA method dataset needs to exist.
- Classification of methods can be used to organise results by sustainability topic.
- Assigning life cycle stages in the eILCD dataset can be used to organise results by life cycle stages.

[4] The OOF can be downloaded from the web tool available at https://orienting.azurewebsites.net/. Section 4.1 in D3.4 (ORIENTING, 2022e) gives more information on how to access the tool.
[5] https://eplca.jrc.ec.europa.eu/LCDN/downloads/ILCD_Format_1.1_Documentation/ILCD_LCIAMethodDataSet.html; accessed 2022-10-06

7. Conclusions

This report describes data specifications for a future LCSA data format following the methodological framework developed in the ORIENTING project (ORIENTING, 2022b) and in relation to the existing eILCD data format (European Commission, 2022): Either existing data fields can be used to meet additional data needs or new fields would need to be added.

The analysis shows that, overall, most additional data needs could one way or the other be satisfied with the existing data structure and data fields. Some additions are needed, mainly a) new flow datasets (having new/other flow types) and b) new flow property datasets (having new/other properties). New flow datasets would be needed for showing other than elementary flows or “artificial” elementary flows as this might be needed for an advanced criticality assessment. The calculation of these flows would need new ways of calculation, but no changes in data format. Furthermore, data points needed for calculation of circularity indicators could also be introduced as exchanges (“other” flows) in process datasets, which also requires new corresponding “artificial” elementary flow datasets. The same is true for social impact scores, which might be considered inventory and impact information at the same time. As inventory information they could be introduced as exchanges and as impact information they could be part of LCIA scores in the LCIA section of process datasets. New flow property datasets would be needed to cover the full list of elements in critical raw materials (if not the entire periodic table) and to cover cost/price/revenue information of products. For additional data to be stored and processed, the “Other” data field in eILCD datasets might be used for many things. However, the preferable solution would be to introduce new data fields with an appropriate name when adapting the eILCD data format to LCSA purposes.

The development of a complete LCSA format would need a similar process as the development of the eILCD data format. The specifications presented in this report can inform such an effort. While the assessment of “artificial” elementary flows would follow the existing structure of LCA, assessing flow properties would be something new and would need discussions also with software providers. One example is S-LCA in ORIENTING, which needs supplier-specific information, which can go beyond the resolution given by activity and location (country or region). Another example is regionalisation where fields for introducing GIS vector files (as evaluated in ORIENTING) or fields for regionalising exchanges are missing.

A consensus data format (likely based on the eILCD format) and the software specifications provided in D3.3 (ORIENTING (2022d)), would expand the potential of integrating full LCSA in commercial software and can be seen as a first step to be considered in an ORIENTING exploitation strategy.

8. References

Benoit-Norris, C., Bennema, M., & Norris, G. (2019). The Social Hotspots Database: Supporting documentation Update 2019 (Vol. 1).

Blengini, G. A., Latunussa, C. E. L., Eynard, U., Torres de Matos, C., Wittmer, D., Georgitzikis, K., Pavel, C., Carrara, S., Mancini, L., Unguru, M., Blagoeva, D., Mathieux, F., & Pennington, D. (2020). Study on the EU's list of Critical Raw Materials (2020) Final Report. https://doi.org/10.2873/904613

Bos, U., Horn, R., & Beck, T. (2016). LANCA ® Characterization Factors for Life Cycle Impact Assessment - Version 2.0. Fraunhofer Verlag.

European Commission. (2021). Commission Recommendation (EU) 2021/2279 of 15 December 2021 on the use of the Environmental Footprint methods to measure and communicate the life cycle environmental performance of products and organisations. Official Journal of the European Union, L 471/1(December 2021), 396. http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=OJ:L:2013:124:TOC

European Commission. (2022). Developer ILCD data format. https://eplca.jrc.ec.europa.eu/LCDN/developerILCDDataFormat.xhtml

Ioannidou, D., Foster, C., Symeonidis, A., Müller, J., Bourgault, G., FitzGerald, D., Moreno Ruiz, E. (2021). Documentation for the 'Allocation, cut-off, EN15804' system model. ecoinvent Association, Zürich, Switzerland.

Maister, K., Di Noi, C., Ciroth, A., & Srocka, M. (2020). PSILCA v.3 - Database documentation.

ORIENTING (2022a). Critical evaluation of social approaches. D1.2 of the ORIENTING project. https://cordis.europa.eu/project/id/958231/results

ORIENTING (2022b). LCSA methodology to be implemented in WP4 demonstrations. D2.3 of the ORIENTING project. https://cordis.europa.eu/project/id/958231/results

ORIENTING (2022c). Documentation of the ORIENTING LCSA ontology (ORIONT). D3.1 of the ORIENTING project. https://cordis.europa.eu/project/id/958231/results

ORIENTING (2022d). Software requirements for user-friendly tools to integrate, visualize and communicate LCSA results. D3.3 of the ORIENTING project. https://cordis.europa.eu/project/id/958231/results

ORIENTING (2022e). Interim User-friendly LCSA tool. D3.4 of the ORIENTING project.

Sonderegger, T. (2021). ISO-adjusted allocation of mineral commodities in ecoinvent. Presentation at LCM 2021 – The 10th International Conference on Life Cycle Management.

Sonderegger, T., Moreno Ruiz, E., Bachmann, T.M., Bianchi, M., Harmens, R., Zamagni, A. (in preparation). Data for LCSA – State-of-the art and perspectives. In: Valdivia, S., Sonnemann, G. (editors). Handbook on Life Cycle Sustainability Assessment.

UNstats (2022a). UNSD – ISIC. https://unstats.un.org/unsd/classifications/Econ/ISIC.cshtml; accessed 2022-10-05

UNstats (2022b). UNSD – CPC. https://unstats.un.org/unsd/classifications/Econ/CPC.cshtml; accessed 2022-10-05

Wernet, G., Bauer, C., Steubing, B., Reinhard, J., Moreno-Ruiz, E., and Weidema, B. (2016). The ecoinvent database version 3 (part I): overview and methodology. The International Journal of Life Cycle Assessment, 21 (9), pp.1218–1230. http://link.springer.com/10.1007/s11367-016-1087-8