universität freiburg

IEDC slide deck

Background material and other information

Status: October 2025

Nildem Atasayar, Stefan Pauliuk, and the ief team

ief Industrial Ecology Freiburg
Research group at the Faculty of
Environment and Natural Resources

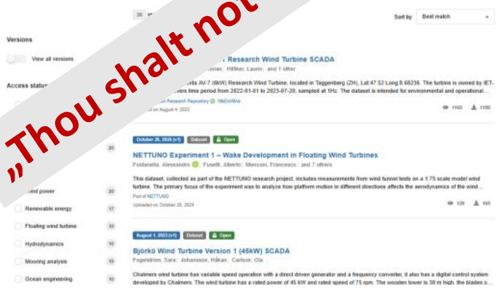
IEDC slide deck
General

Types of databases and their business cases

Data lakes (Zenodo)

- broad scope
- co-existence

„Thou shalt not be forgotten“



missing link

iedc

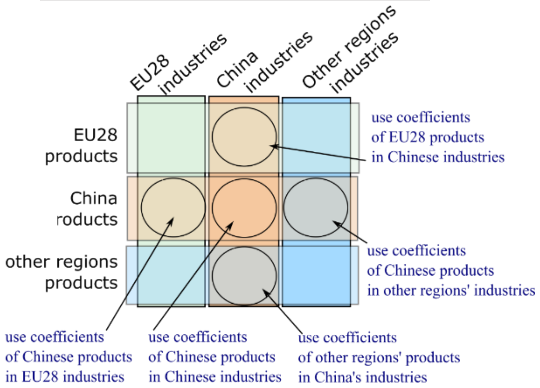
Integrated databases (e.g., ecoinvent, MRIO)

- systematic structure
- consistent labels

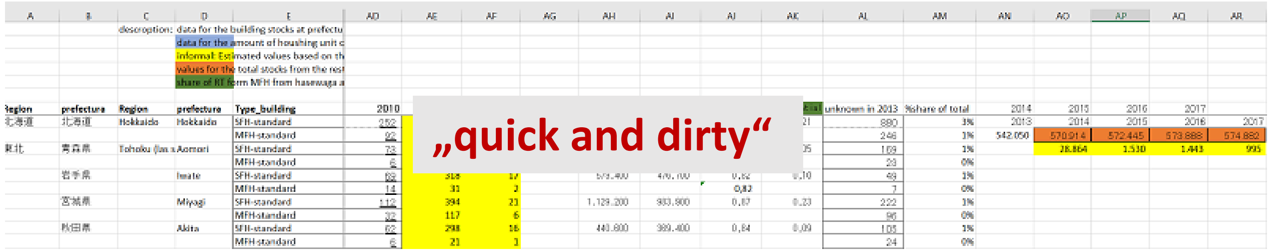
„right tool to get the job done“

- small data
- spreadsheet workflow

- link data to the phenomena they describe
- data are not linked to tools but interoperable across tools
- open access and transparency

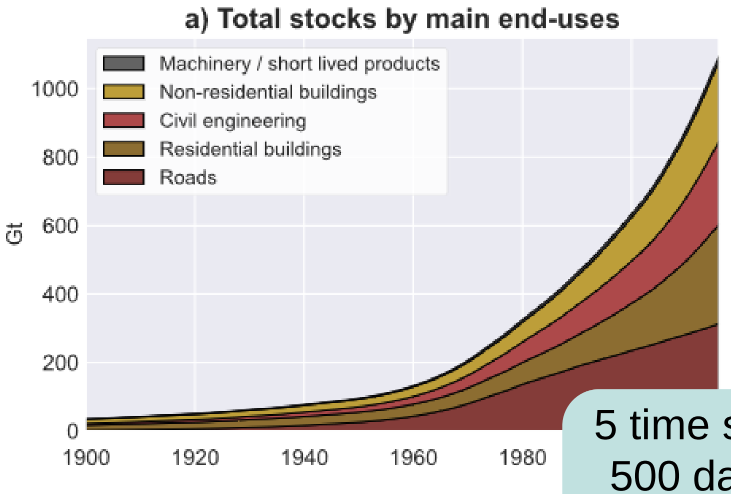


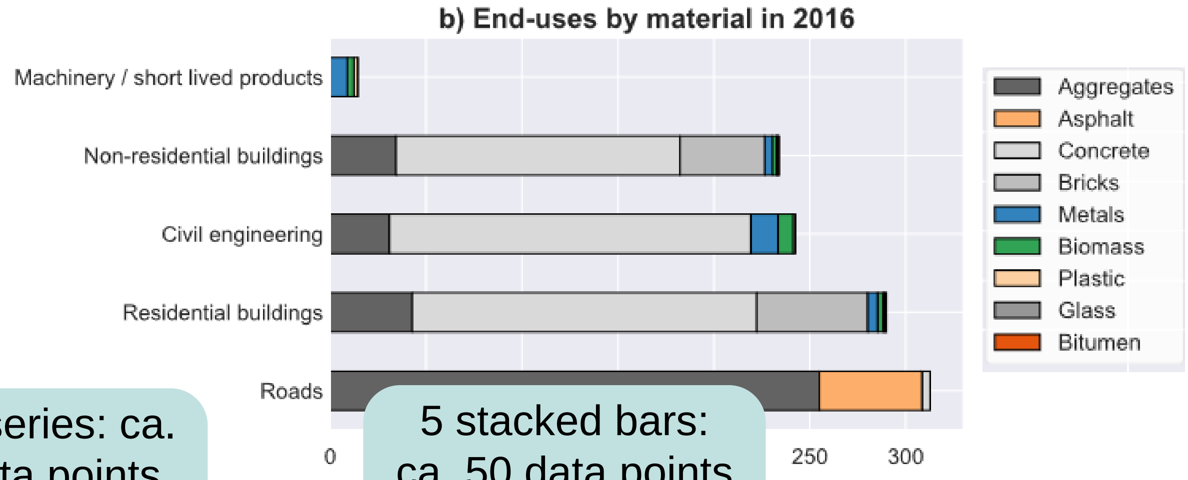
‚The usual mess‘



„quick and dirty“

Data points and datasets in the IEDC, examples





5 time series: ca. 500 data points in a dataset for an in-use stock

5 stacked bars: ca. 50 data points in a dataset for a sector split

3 process steps: ca. 20 data points in a dataset for market shares

		Process	% share	Process	% share
		Continuous casting	69	Section mill	37
Coking	3.7	Ingot casting	49	Cold rolling	46
Blast furnace	1.8	Steel casting	65	Galvanizing	61
Direct reduction	5.2	Iron casting	67	Tinning	72
Scrap preparation	100	Primary mill	37	Forming	37
BOF	8.3	Hot rolling	36	Extrusion	64
EAF/DRI	77	Plate mill	37	Coating	94
EAF	83	Rod and bar mill	29		

lists of production, consumption, or trade flows: 10000-50000 data points per dataset

material	commodity				unit
steel	passenger vehicle, intern			2775	kg
cast iron	passenger vehicle, intern			2446	kg
aluminium	passenger vehicle, intern			9941	kg
copper	passenger vehicle, interna			36603	kg
magnesium	passenger vehicle, internal combu	Global	2000-2010	0.240403956	kg
zinc	passenger vehicle, internal combu	Global	2000-2010	0.099790321	kg
lead	passenger vehicle, internal combu	Global	2000-2010	11.54087636	kg
gold	passenger vehicle, internal combu	Global	2000-2010	0	kg
silver	passenger vehicle, internal combu	Global	2000-2010	0	kg
neodymium	passenger vehicle, internal combu	Global	2000-2010	0	kg
platinum	passenger vehicle, internal combu	Global	2000-2010	0.0011	kg
paladium	passenger vehicle, internal combu	Global	2000-2010	0.0053	kg
rhodium	passenger vehicle, internal combu	Global	2000-2010	0.0006	kg
arsenic	passenger vehicle, internal combu	Global	2000-2010	0.004940321	kg
antimony	passenger vehicle, internal combu	Global	2000-2010	0.116920923	kg
plastic	passenger vehicle, internal combu	Global	2000-2010	186.3873028	kg

In the IEDC, datasets are assemblies or collections of individual data points, defined by a common data type and by a common context.

Each IEDC dataset has the same structure

Dataset description

- dataset name & ID
- dataset grouping & type
- dataset scope & resolution
- dataset description
- dataset aspects and classifications (semantic metadata)
- dataset source and license
- dataset submission and review
- reserved tags

Dataset information		
	Column name	**Dataset entries**
Identification	dataset_id	auto
	dataset_name	3_MC_Vehicles_Hawkins_2012
	dataset_version	NULL
Grouping	datagroup_id	NULL
	data_category	3
	data_type	Material composition
	data_layer	Mass per unit
System location: What elements and objects in the system are described?	process_scope	none
	process_resolution	none
	product_scope	passenger vehicles
	product_resolution	3 vehicle types: ICE vehicle, battery electric vehicl
	material_scope	metals, polymers, silica material, chemicals, other
	material_resolution	42 materials
	regional_scope	global
	regional_resolution	global
	temporal_scope	ca. 2000-2010
	temporal_resolution	ca. 2000-2010
Description	description	Material composition of passenger vehicles, extrac
	keywords	product material composition; metal content; pass
	data_provenance	Industry data
	dataset_size	0
	comment	none
System location: Dataset aspects and semantics. For aspects: [Aspect symbol; classification]. Mandatory aspects in gray shading.	aspect_1	engineering_material
	aspect_1_classification	4
	aspect_2	commodity
	aspect_2_classification	7
	aspect_3	region
	aspect_3_classification	2
	aspect_4	age-cohort
	aspect_4_classification	14
	aspect_5	none
	aspect_5_classification	none
	aspect_6	none
	aspect_6_classification	none
	aspect_7	none
	aspect_7_classification	none
	aspect_8	none
	aspect_8_classification	none
	aspect_9	none
	aspect_9_classification	none
	aspect_10	none
	aspect_10_classification	none
	aspect_11	none
	aspect_11_classification	none
	aspect_12	none
	aspect_12_classification	none
	tupel_notation	Value(m,g,r,c)
	semantic_string_exampl	The typical global zinc content of ICE vehicles, ca. 2
	semantic_string_genera	[layer] for [data type] of [aspect1] in [aspect2] in [
Data access and licence	type_of_source	Proprietary report or research article
	project_license	all rights reserved
	main_author	Troy R. HAWKINS
	dataset_link	https://onlinelibrary.wiley.com/doi/abs/10.1111/j.15
	dataset_format	Excel spreadsheet
	project_report	DOI: 10.1111/j.1530-9290.2012.00532.x
	suggested_citation	DOI: 10.1111/j.1530-9290.2012.00532.x
	visible	1
Data submission, review, and conversion info	access_date	42281
	submission_date	2018-06-04
	submitting_user	Stefan Pauliuk
	dataset_conversion_info	Numbers converted from Excel to database via Pytl
	review_date	
	review_user	
	review_comment	
Reserve	reserve1	
	reserve2	
	reserve3	
	reserve4	
	reserve5	

data points of dataset: list or table

List-shaped data, with unit, uncertainty, and comment for each data point

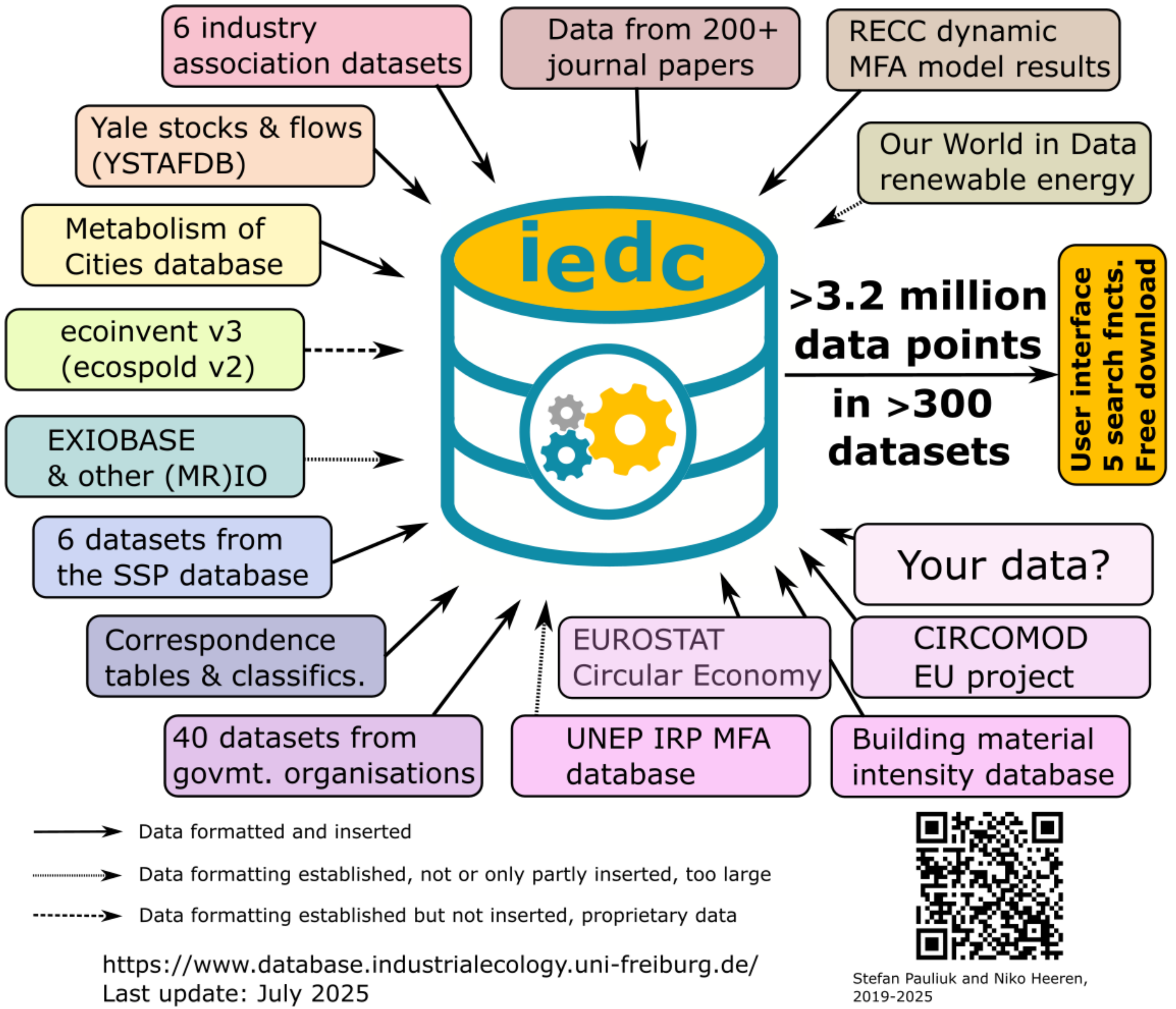
material	commodity	region	age-cohort	value	unit
steel	passenger vehicle, internal combu	Global	2000-2010	882.0902775	kg
cast iron	passenger vehicle, internal combu	Global	2000-2010	115.8742446	kg
aluminium	passenger vehicle, internal combu	Global	2000-2010	71.2499941	kg
copper	passenger vehicle, internal combu	Global	2000-2010	23.87236603	kg
magnesiu				.240403956	kg
zinc				.099790321	kg
lead				1.54087636	kg
gold				0	kg
silver				0	kg
neodymium	passenger vehicle, internal combu	Global	2000-2010	0	kg
platinum	passenger vehicle, internal combu	Global	2000-2010	0.0011	kg
paladium	passenger vehicle, internal combu	Global	2000-2010	0.0053	kg
rhodium	passenger vehicle, internal combu	Global	2000-2010	0.0006	kg
arsenic	passenger vehicle, internal combu	Global	2000-2010	0.004940321	kg
antimony	passenger vehicle, internal combu	Global	2000-2010	0.116920923	kg
plastic	passenger vehicle, internal combu	Global	2000-2010	186.3873028	kg

Table-shaped data, with any number of row and column aspects

				residential buildings (all	single-family house	multi-family houses	apartment blocks
use phase - resid	Austria	2015	before 1945	78,7339048	44,5912	15,9589	18,1838048
use phase - resid	Austria	2015	1945-1969	92,2876096	60,5438048	14,39	17,3538048
use phase - resid	Austria	2015	1970-1979	53,3176096	39,0938048	5,8	8,4238048
use phase - resid	Austria					8	3,8238048
use phase - resid	Austria					3	5,0738048
use phase - resid	Austria					8	4,78
use phase - resid	Austria					4	1,79
use phase - resid	Belgium					5	3,695
use phase - resid	Belgium	2015	1945-1969	63,119375	4,255	33,864375	25
use phase - resid	Belgium	2015	1970-1979	69,22972655	1,48972655	34,2325	33,5075
use phase - resid	Belgium	2015	1980-1989	57,8744531	0,91972655	23,86722655	33,0875
use phase - resid	Belgium	2015	1990-1999	60,20613675	3,07972655	32,6414102	24,485

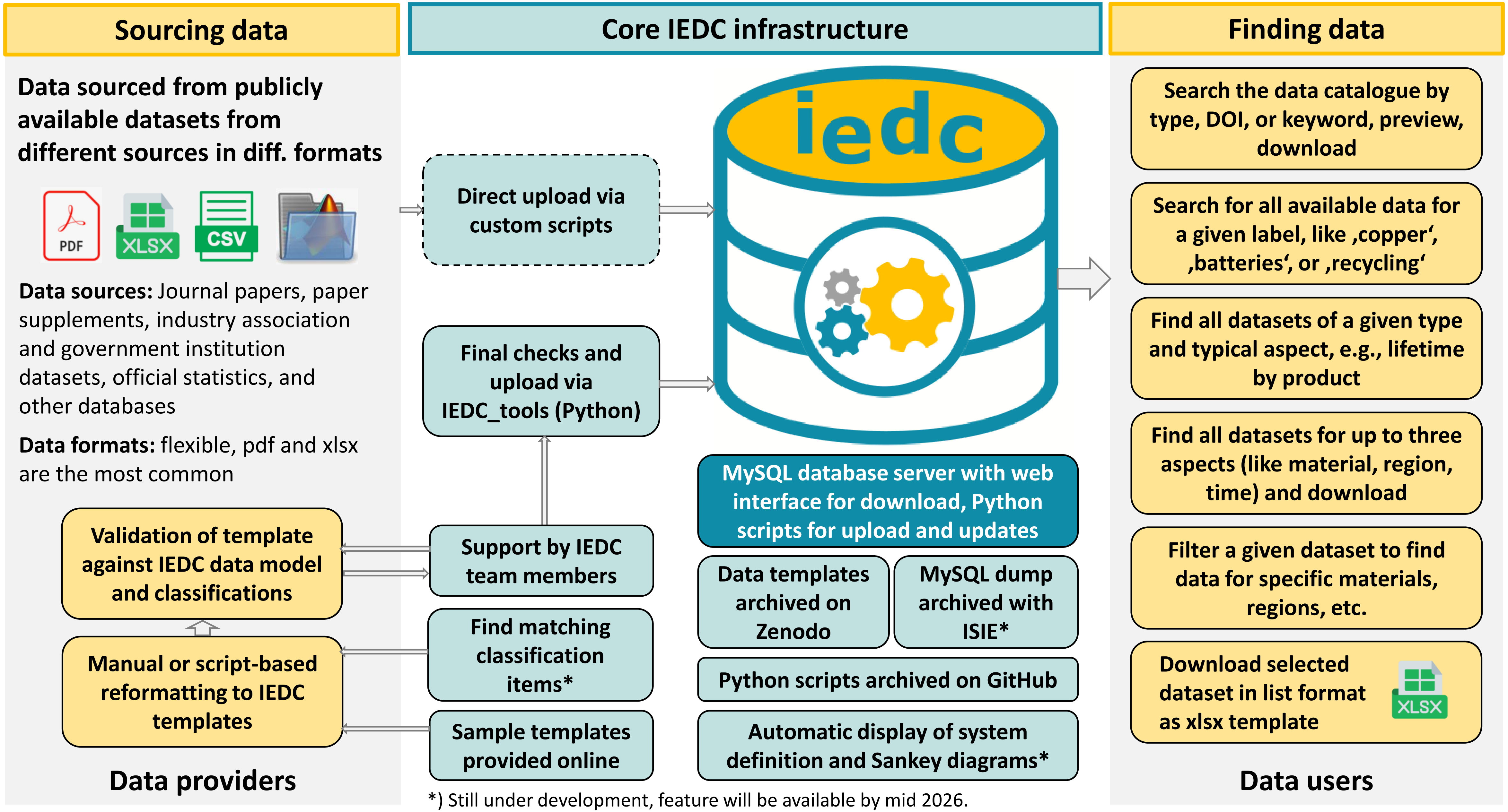
Source: „A General Data Model for Socioeconomic Metabolism and its Implementation in an Industrial Ecology Data Commons Prototype ", by Pauliuk et al. (2019), JIE, DOI: 10.1111/jiec.12890

Industrial ecology data commons – IEDC



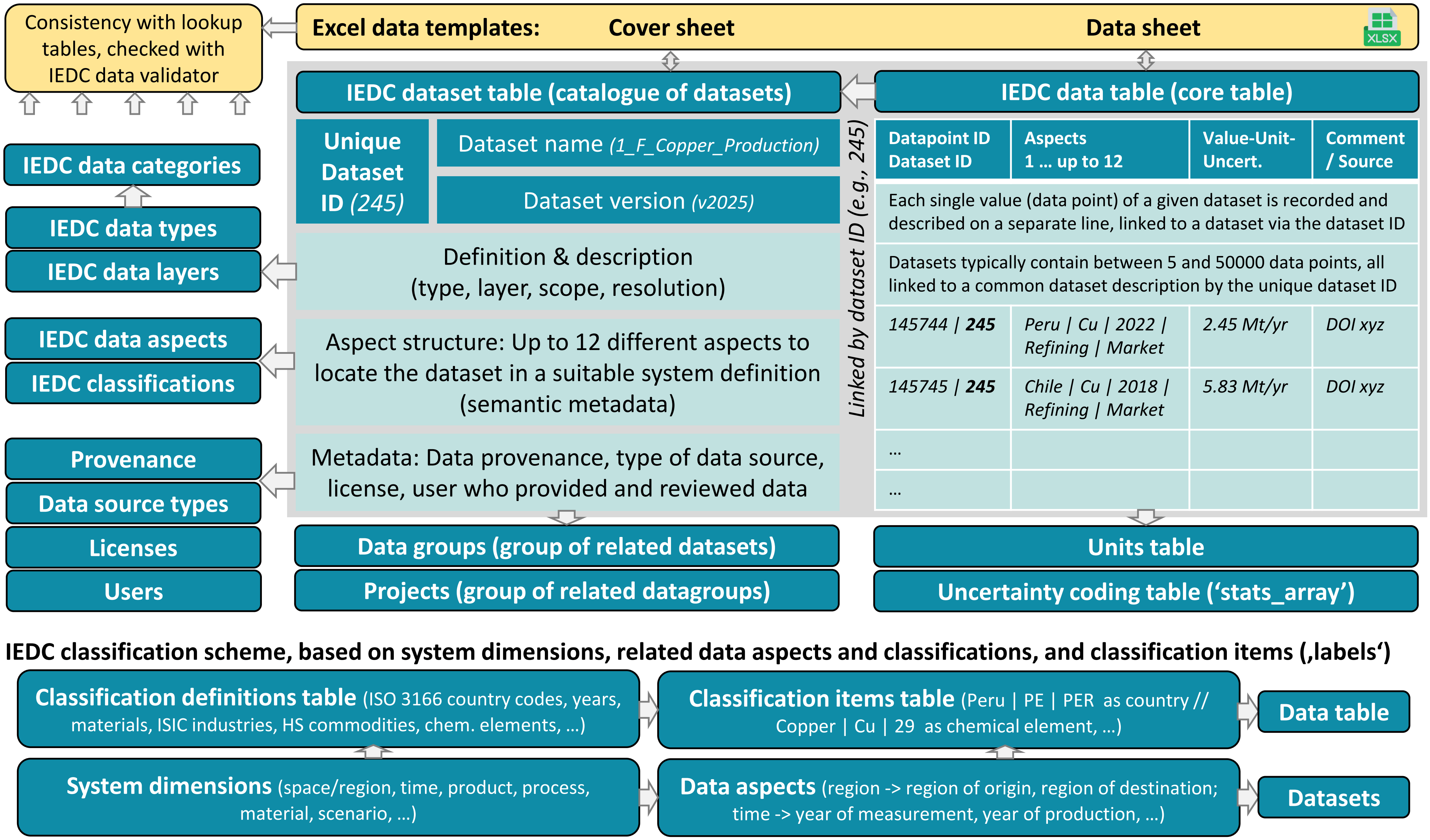
https://www.database.industrialecology.uni-freiburg.de/

IEDC infrastructure, services, and data workflow



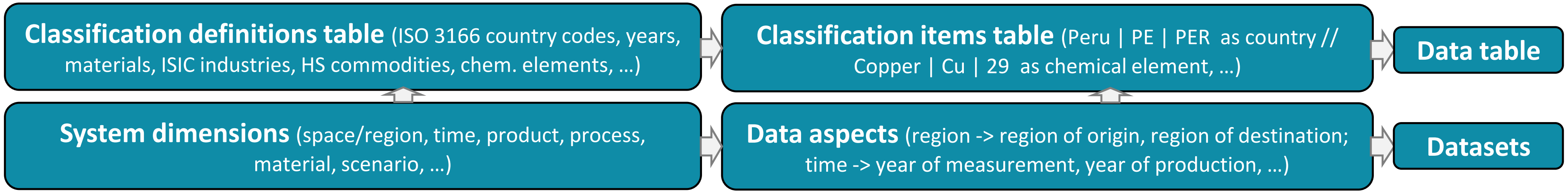
*) Still under development, feature will be available by mid 2026.

Structure of the IEDC database (relations between IEDC tables)



IEDC classifications: A large and comprehensive compilation of general classifications for socio-metabolic data

IEDC classification scheme, based on system dimensions, related data aspects and classifications, and classification items (‚labels‘)



About 100 classifications are available for consistent data labelling:

- ISO regions, chemical elements,
- materials
- products and components
- trade codes
- industry codes and industries/processes
- LCI compartments and layers

All classifications can be browsed and downloaded via

https://www.database.industrialecology.uni-freiburg.de/classifications.aspx

UNI FREIBURG

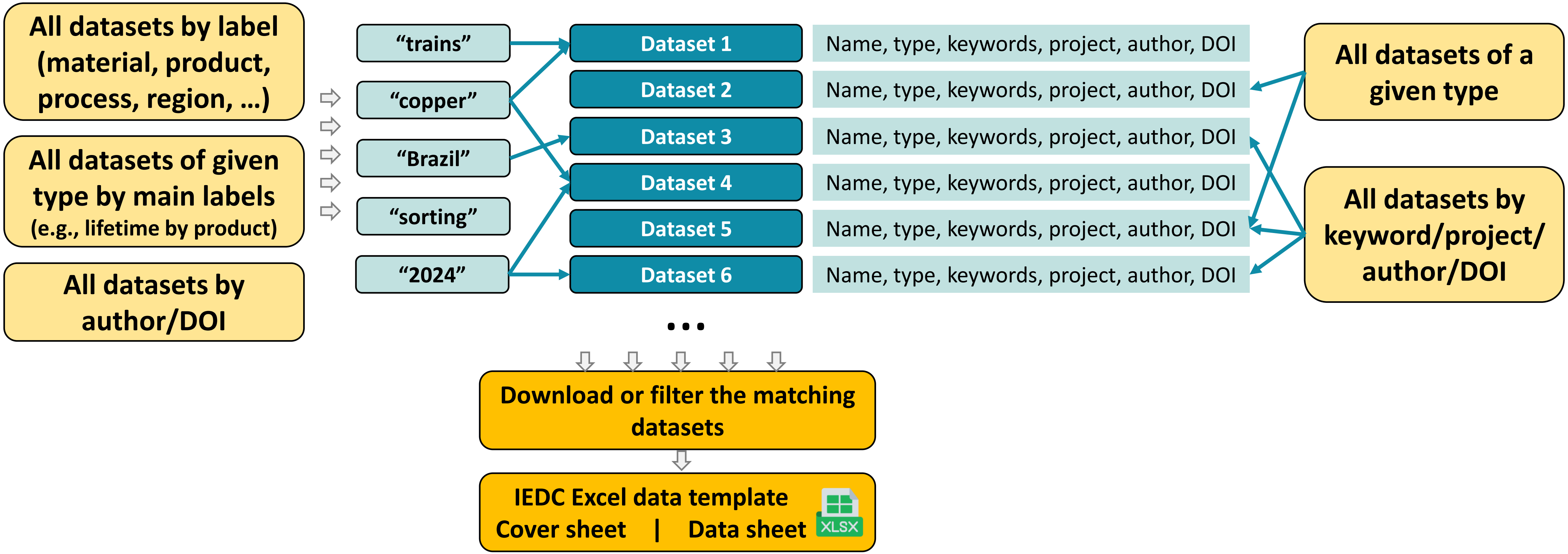
Core IEDC vocabulary

- **Data point or data item:** Individual number with unit and uncertainty (optional), that occupies a single and unique row in the IEDC‘s data table. Each data point is part of exactly one dataset.
- **Dataset:** Assembly or collection of individual data points, defined by a common data type AND by a common context, e.g., data from a certain publication or on a certain sector, material, or topic, like steel production scenarios, in-use stock of vehicles, etc. Each dataset is described on one unique row in the IEDC‘s dataset table.
- **Datatype:** A certain kind or species of data with the same distinct meaning and location in system: e.g., a flow of something from some process to some other process, lifetime of something in some process; unit process inventory of a certain activity.

- **Aspects:** To explain the exact the meaning of a dataset, the IEDC allows for adding up to 12 attributes that describe the „what“, „when“, and „where“ of the data (semantic metadata). These attributes are always relative to a system definition that is specific to each dataset, and they form aspects in that system, such that ‘what‘ can refer to the aspects commodity, material, or chemical element; ‘when‘ can refer to the aspects year of measurement or year of production; and ‘where‘ can refer to the aspects ‘in a process, ‘into a process‘ or ‘out of a process‘.
- **Dimensions, classifications and classification items:** The different aspects describe how the different datasets relate to the system dimensions (time, space/region, process, material). These dimensions are measured in discrete units (e.g., years, countries, industrial sectors, chemical elements, or engineering materials), and a list of such labels for the discrete units is called classification (e.g., ISO 3166 country codes), and a member of this list a classifiation item (e.g., ‘Brazil‘). All aspects of all datasets use a predefined classification item from an IEDC classification.

Search perspectives in the IEDC

All data in the IEDC are assembled into datasets, and there are two types of searches: One that searches the data catalogue for a data types, keywords, projects, author names, or DOIs (right), and one that searchers the aspects and the comments (that may contain autor names and DOIs) of all individual data points (left).



Overview of the different IEDC search options

I am looking for…

I am looking for…	Search option
… all datasets that match a given keyword … all datasets of a certain data type … all datasets from a certain project … a dataset with a given name or id	**Data catalogue search by keyword (see bottom of this page)**
… all datasets that contain a certain product/material/region e.g., what data are there on copper / Brazil ?	**Find all datasets with a given label (see link via button below)**
… all datasets of a given type that contain a certain product/material/region e.g., are there any lifetime data for bicyles?	**Search for datasets by main aspects (see link via button below)**
… all datasets of a given type that contain any combination of labels (material, product, process) for up to three different aspects	**Search for datasets by all aspects (see link via button below)**
…data for a specific label (material, region, product, process, …) in a given dataset	**Filter a dataset for a given label (see link via button below)**
…data from a specific publication (DOI) or from a certain author	**Search for datasets by DOI or author (see link via button below)**

Download button on the different search pages

IEDC Excel data template
Cover sheet | Data sheet
XLSX

Typical IEDC Use Cases

LCA – quick access to foreground data
material composition of products, specific energy requirement of product operation, product lifetime, manufacturing yield, process parameters

MFA – stocks, flows, and parameters
Large library of MFA study results for stocks and flows of products and materials, MFA system parameters such as material composition of products, EoL-recovery rates

Policy analyses
Data on material criticality, environmental pressure indicators

Urban metabolism studies
City-level stocks, flows, and impacts

Scenario modelling
Data on service provision, in-use stocks, alternative technology process inventories, potentials for resource efficiency in manufacturing and recycling

https://tinyurl.com/iedc-freiburg

IEDC slide deck

Data model

A General Data Model for Socioeconomic Metabolism and its Implementation in an Industrial Ecology Data Commons Prototype

Stefan Pauliuk, Niko Heeren, Mohammad Mahadi Hasan, and Daniel B Müller

UNI FREIBURG

What the IEDC is and what it is not

The IEDC is a database of quantitative descriptions of socio-ecomic metabolism (society‘s use of energy and materials), ranging from data on individual products and materials to society-wide statistics on stocks and flows.

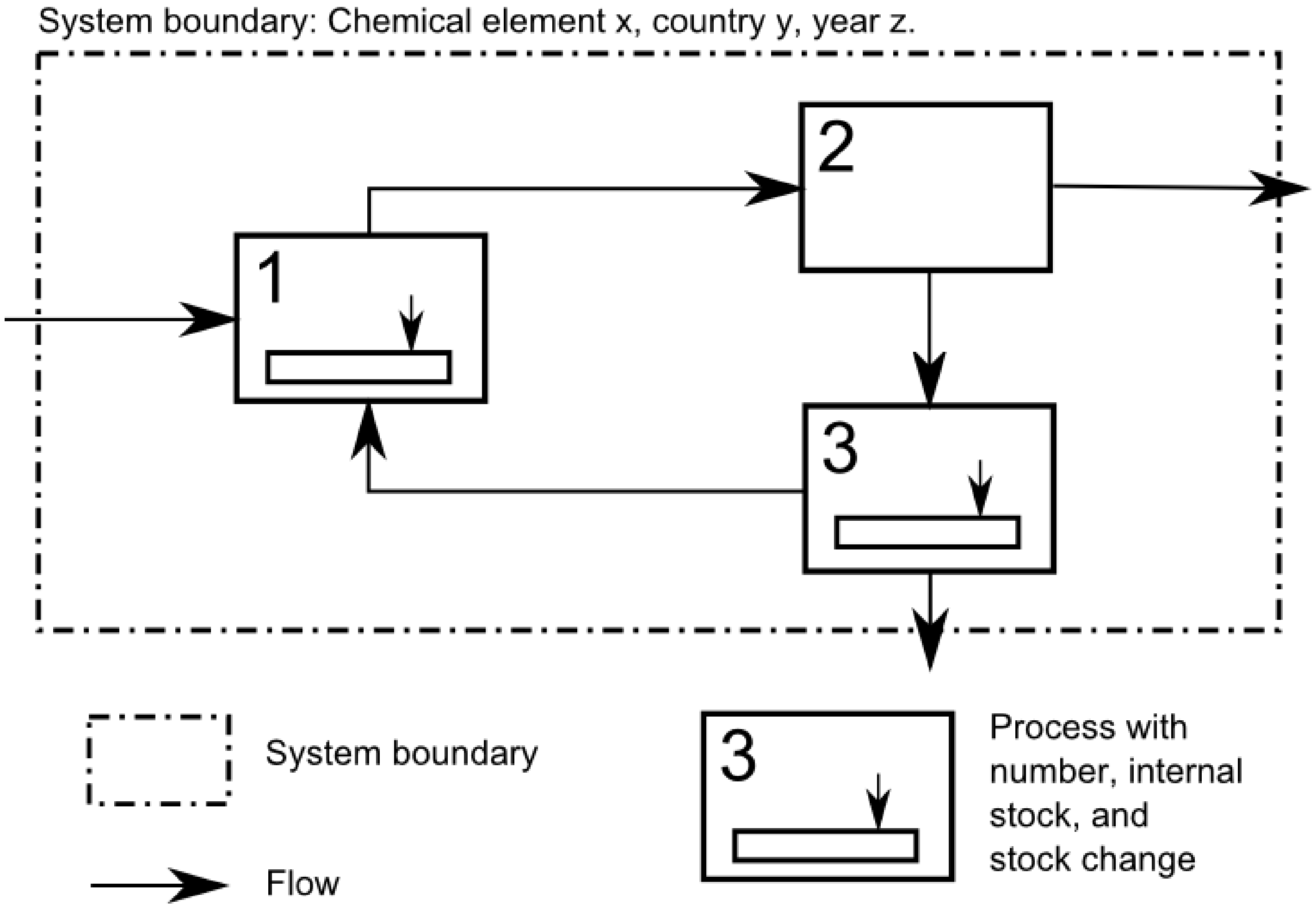
IEDC data inform industrial ecology and socio-metabolic research, including MFA, LCA, IO, econometric, and other quantitative studies.

IEDC is a comprehensive database with an intermediate level of data integration:

Integration level	Semantics/ structure	Aspect classification	Accessibility	Data linking
Low, e.g. Zenodo	No data model used	Custom classifications prevail	All data formats are allowed, including figures and pdfs.	No linkages, possibly shared classifications without harmonized identifiers.
Intermediate, e.g. IEDC (this work)	Common data model used	Mix of general and custom classifications	All datasets are machine-readable in common format	Thematic linkages, some shared classifications with harmonized and unique identifiers.
High, e.g. ecoinvent	Common semantic model across all datasets	Common classification used across all datasets	All datasets are machine-readable in common format	Fully linked datasets with shared classification and unique identifiers

Source: „A General Data Model for Socioeconomic Metabolism and its Implementation in an Industrial Ecology Data Commons Prototype “, by Pauliuk et al. (2019), JIE, DOI: 10.1111/jiec.12890

The basic elements of a system definition



System boundary:
Shows what processes are part of the system studied. System needs to be specified in space and time.

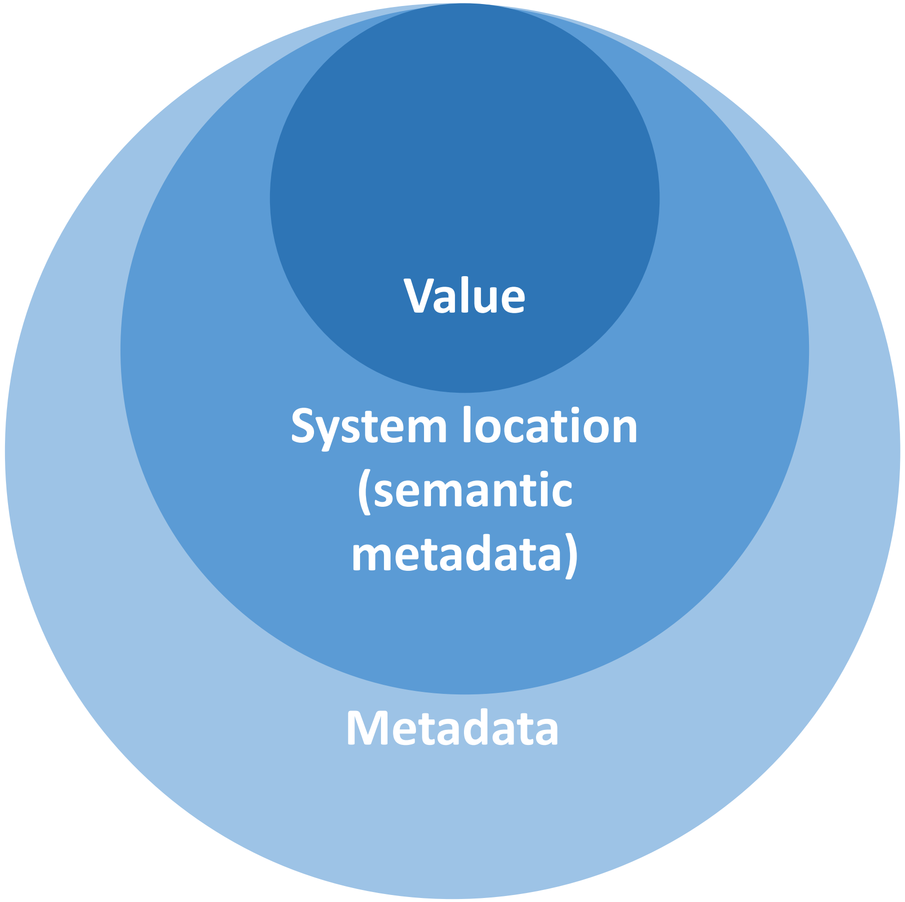
Process: Element of a system were material or energy is transformed, stored, or distributed.

Stock: Storage of material, products, or energy. A stock is always associated with a process. It is measured at a given point of time t. ‚snapshot‘

Flow: Transport of material products, or energy across the system. A flow always connects two processes or one process with the environment. It is measured over an interval $[t_1,t_2]$.

Data in a systems context

Quantitative information on processes, stocks, flows, etc., in socioeconomic metabolism has three components:



1) **Value.** The actual numerical information, including unit and uncertainty
2) **System location (semantic metadata).** The information needed to locate the data in the systems context, i.e., the link between data and the system dimensions (process, time, region, material, …)
3) **Metadata.** Information like provenance, source document, author and version information, and license.

Source: „A General Data Model for Socioeconomic Metabolism and its Implementation in an Industrial Ecology Data Commons Prototype ", by Pauliuk et al. (2019), JIE, DOI: 10.1111/jiec.12890

Data aspects and system dimensions I

- **System definition prescribes a number of *dimensions* along which the system content is described:**

- the *time dimension* is used to order events by the time of their occurrence
- the *location dimension* is used to order objects by their location
- the *process dimension* is used to identify balance volumes or to group events
- the *object dimension* is used to identify different goods or substances
- the *layer dimension* is used to indicate the unit in which the data are measured

Source: „A General Data Model for Socioeconomic Metabolism and its Implementation in an Industrial Ecology Data Commons Prototype ", by Pauliuk et al. (2019), JIE, DOI: 10.1111/jiec.12890

UNI FREIBURG

Data aspects and system dimensions II

To locate data in a system definition, one has to specify the *aspects* of the data that describe how the given values relate to time, region, processes, etc. dimensions in the system.

The aspects describe how the system dimensions relate to the data.

For example, a flow has a starting node and a terminating node. Here, the system dimension 'process' is used to describe two aspects of a flow, namely the starting and terminating process node.

A stock is always associated with a node where it is located, and therefore, 'residence process' is an aspect of the 'process' dimension needed to locate a stocks

Source: „A General Data Model for Socioeconomic Metabolism and its Implementation in an Industrial Ecology Data Commons Prototype ", by Pauliuk et al. (2019), JIE, DOI: 10.1111/jiec.12890

IEDC data categories and data types

	Objects of interest (goods, products, substances, commodities, waste, …)	Processes (industries, markets, end-use sectors, use-phase)
Extensive (at scale)	**Flows (1)** • Flow • Process inventory • Unit process inventory • Births/Deaths **Stocks (2)** • Stock • In-use stock • Population	**Extensive process properties (5)** • Process output capacity • Employment (work places)
Intensive (per unit)	**Intensive object properties (3)** • Product material composition • Product lifetime • Price of products • Specific energy consumption	**Intensive process properties (4)** • Process yield factors • Process environmental extensions per output • Process operating costs per output
Other data	**General ratios (6)** • Per capita stock • Per capita flow • Material substitution coefficient **Correspondence tables (7)**	

Source: „A General Data Model for Socioeconomic Metabolism and its Implementation in an Industrial Ecology Data Commons Prototype ", by Pauliuk et al. (2019), JIE, DOI: 10.1111/jiec.12890

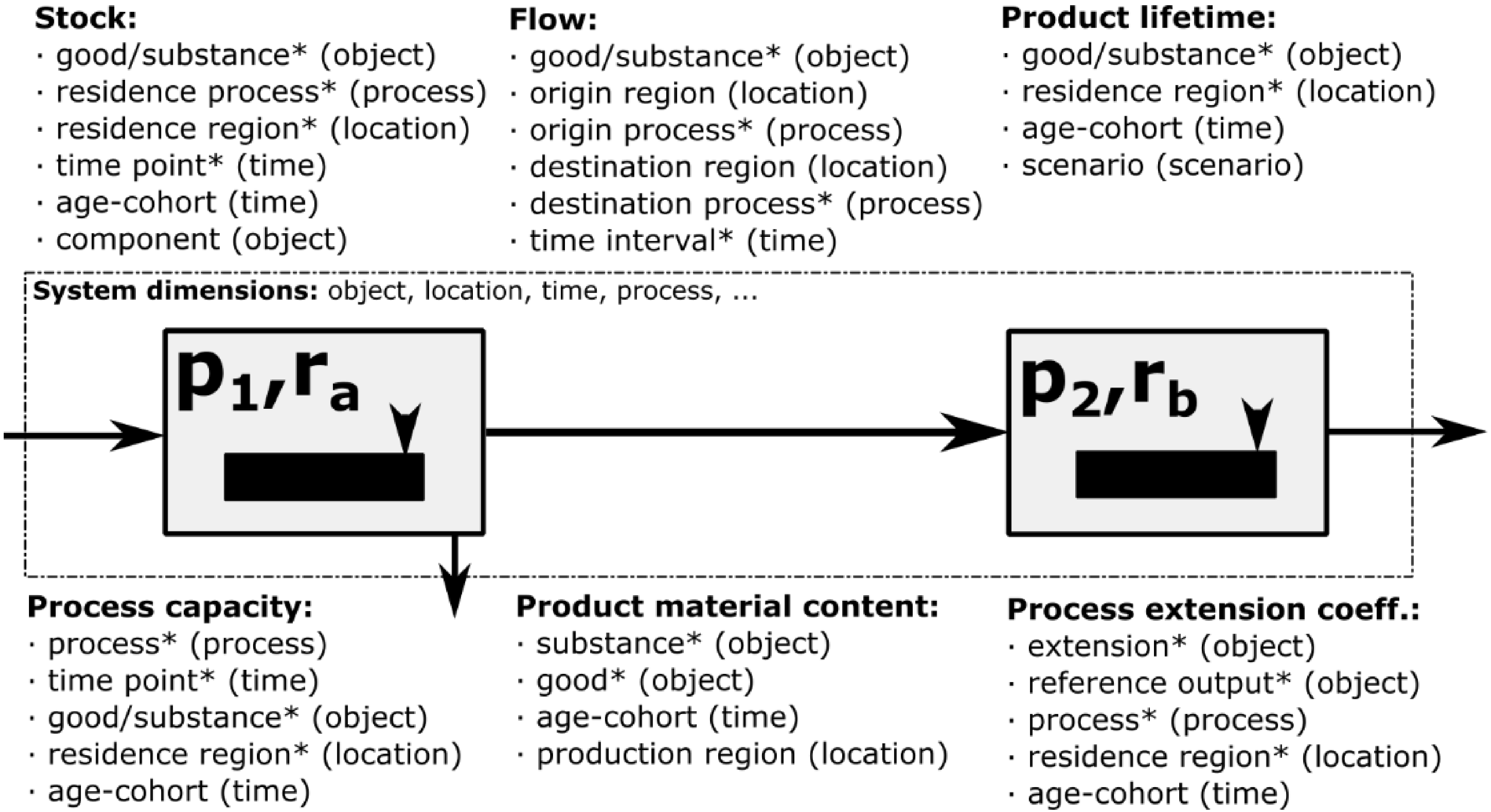
Data aspects and system dimensions (examples)

System dimension	Description	Related data aspects (example)
Layer	Unit of measurement	Dry mass, energy, number of items, volume, economic value
Process	Transformation, distribution, storage events	Process of residence (stock), process of origin (flow), process of destination (flow)
Location	Location in space	Region of residence (stock), region of origin (flow), region of destination (flow)
Object	Objects of interest (goods, substances, commodities, waste, products, …)	Commodity, good, product group, product type (sub-product), substance, chemical element, waste type, environmental extension
Time	Location in time	Historic time, model time, age-cohort, time point (stock), time interval (flow)
Scenario	Describing different "realities" or manifestations	Scenario for model drivers, scenario for process parameters

Table: The common system dimensions for socioeconomic metabolism
and the related data aspects. The list of dimensions and aspects is not exhaustive.

Source: „A General Data Model for Socioeconomic Metabolism and its Implementation in an Industrial Ecology Data Commons Prototype ", by Pauliuk et al. (2019), JIE, DOI: 10.1111/jiec.12890

Data types (examples)



Source: „A General Data Model for Socioeconomic Metabolism and its Implementation in an Industrial Ecology Data Commons Prototype ", by Pauliuk et al. (2019), JIE, DOI: 10.1111/jiec.12890

Data aspects and system dimensions VI

(D1) Each data item (number quantifying a fact in a system) requires a minimum number of aspects to be meaningfully located in the system dimensions.

(D2) Each data type (stock, flow, material content, product lifetime, …) has a specific data model that prescribes which aspects are required and which aspects are optional for the meaningful location of this data type in the system definition.

Source: „A General Data Model for Socioeconomic Metabolism and its Implementation in an Industrial Ecology Data Commons Prototype ", by Pauliuk et al. (2019), JIE, DOI: 10.1111/jiec.12890

Examples I

- **Flows (category 1):** A flow is an extensive system variable describing the relocation of objects: [Flow] of [objects] from [process A] in [region of origin] to [process B] in [region of destination] in [time period] is [value/uncertainty] for [layer].

- **Stocks (category 2):** A stock is an extensive system variable describing the location of objects: [Stock] of [objects] of [age-cohort (*)] in [process] in [region of location] at [time point] is [value/uncertainty] for [layer].

Source: „A General Data Model for Socioeconomic Metabolism and its Implementation in an Industrial Ecology Data Commons Prototype ", by Pauliuk et al. (2019), JIE, DOI: 10.1111/jiec.12890

UNI FREIBURG

Examples II

- **Product lifetime (category 3):** The lifetime is an intensive object property describing the residence time of a good/substance in a process: [Lifetime] of [good/substance] of [age-cohort (*)] in [process] in [region] is [value/uncertainty] for [layer].

- **Process yield factors (category 4):** The process yield factor is an intensive process property describing the share of a material in an input good/substance that is transformed or manufactured into an output good/substance: [Yield] of [material] in [input good/substance] into [output good/substance] in [process] of [technology (*)] of [age-cohort (*)] in [region] is [value/uncertainty] for [layer]. The layer can be mass but also volume.

Source: „A General Data Model for Socioeconomic Metabolism and its Implementation in an Industrial Ecology Data Commons Prototype ", by Pauliuk et al. (2019), JIE, DOI: 10.1111/jiec.12890

Examples III

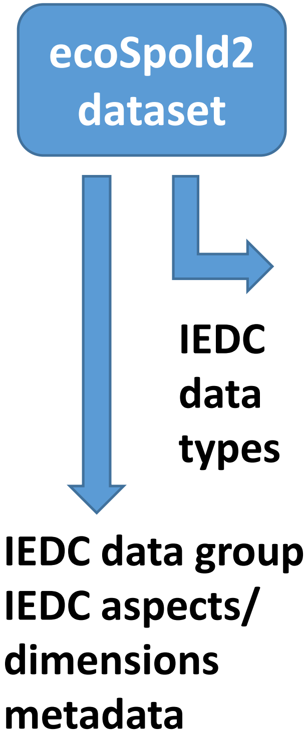
- **Process capacity (category 5):** The process capacity is an extensive process property describing the maximum rate of output commodity that can be produced in a certain process: [Capacity] of producing [good/substance] in [process] of [technology (*)] of [age-cohort (*)] in [region] is [value/uncertainty] for [layer].

- **Per capita stock (category 6):** The per capita stock is a general ratio describing the amount of a commodity per person in a certain process: [Per capita stock] of [good/substance] of [age-cohort (*)] in [process] in [region] is [value/uncertainty] for [layer]. The layer can be mass, number of items, or monetary value.

Source: „A General Data Model for Socioeconomic Metabolism and its Implementation in an Industrial Ecology Data Commons Prototype ", by Pauliuk et al. (2019), JIE, DOI: 10.1111/jiec.12890

Example IV: mapping ecoSpold2 to IEDC data types

Why do we need this? Doesn't LCA already have a comprehensive data model in form of *ecospold*?
Ecospold contains activity/process inventories (exchanges) and related data: price and production volumes of reference flows and the composition (carbon content, water content) of substances.
But there are many more data that are needed in industrial ecology research, including: product lifetimes, breakdown of products into components and materials, process capacities, population, economic statistics, trade flows, inequality data, or economy-wide and sector-specific material flow accounts.



List of datasets created, either for each activity separate or for database as a whole:

Aspect numbers	1	2	3	4	5	6	7
1_F_ProductionVolumes_<database name	**material**	**origin_proces**	**destination_**	**origin_region**	**destination_**	**time**	**layer**
Production volumes	product/material Id (ecospoldv2: exchar	activity id	unspecified	geographical :	unspecified	time interval for which volume	mass, energy
6_FPR_UnitPrices_<database name>	**material**	**origin_proces**	**destination_**	**origin_region**	**destination_**	**time**	**layer**
Unit prices for exchanges	product/material Id (ecospoldv2: exchar	activity id	unspecified	geographical :	unspecified	time frame of dataset.	price
1_PI_<database name>_<activity id> or 1_	**material**	**material_cate**	**origin_proces**	**destination_**	**origin_region**	**destination_region**	**time**
Process inventory (flows)	product/material Id (ecospoldv2: exchar	Elementary or	for inflows: ac	for outflows: ;	for inflows: nc	for outflows: none. For inflows:	time frame of dataset. F
6_MIP_<database name>_<activity id>	**material**	**origin_proces**	**destination_**	**origin_region**	**destination_**	**time**	**layer**
Intensive flow properties other than price	product/material Id (ecospoldv2: exchar	activity id	unspecified	geographical :	unspecified	time frame of dataset.	carbon content, dry mat
4_PAR_<database name>_<activity id>	**process**	**region**	**time**	**layer**			
Parameters	activity Id	geographical :	time frame of	parameter name			
4_TC_<database name>_<activity id>	**process**	**region**	**time**	**input_materi**	**output_mate**	**layer_in**	**layer_out**
Transfer coefficients	activity Id	geographical :	time frame of	@ExchangeId	product/mate	mass, energy, ...	mass, energy, ...
6_IMI_ImpactIndicators_<database name:	**material**	**origin_proces**	**destination_**	**region**	**time**	**layer**	**impact_indicator**
Impact indicator results	product/material Id (ecospoldv2: exchar	activity Id	compartment	geographical :	time frame of	mass, energy	Reference to the impact

https://github.com/IndEcol/IE_data_commons/blob/master/doc/LCI_IEDC_Correspondences_V1.xlsx

Using the data model I

dataset table entry

Dataset information		
	Column name	Dataset entries
Identification	dataset_id	auto
	dataset_name	3_MC_Vehicles_Hawkins_2012
	dataset_version	NULL
Grouping	datagroup_id	NULL
	data_category	3
	data_type	Material composition
	data_layer	Mass per unit
System location: What elements and objects in the system are described?	process_scope	none
	process_resolution	none
	product_scope	passenger vehicles
	product_resolution	3 vehicle types: ICE vehicle, battery electric vehicl
	material_scope	metals, polymers, silica material, chemicals, other
	material_resolution	42 materials
	regional_scope	global
	regional_resolution	global
	temporal_scope	ca. 2000-2010
	temporal_resolution	ca. 2000-2010
Description	description	Material composition of passenger vehicles, extrac
	keywords	product material composition; metal content; passe
	data_provenance	Industry data
	dataset_size	0
	comment	none
System location: Dataset aspects and semantics. For aspects: [Aspect symbol; classification]. Mandatory aspects in gray shading.	aspect_1	engineering_material
	aspect_1_classification	4
	aspect_2	commodity
	aspect_2_classification	7
	aspect_3	region
	aspect_3_classification	2
	aspect_4	age-cohort
	aspect_4_classification	14
	aspect_5	none
	aspect_5_classification	none
	aspect_6	none
	aspect_6_classification	none
	aspect_7	none
	aspect_7_classification	none
	aspect_8	none
	aspect_8_classification	none
	aspect_9	none
	aspect_9_classification	none
	aspect_10	none
	aspect_10_classification	none
	aspect_11	none
	aspect_11_classification	none
	aspect_12	none
	aspect_12_classification	none
	tupel_notation	Value(m,g,r,c)
	semantic_string_exampl	The typical global zinc content of ICE vehicles, ca. 2
	semantic_string_genera	[layer] for [data type] of [aspect1] in [aspect2] in [
Data access and licence	type_of_source	Proprietary report or research article
	project_license	all rights reserved
	main_author	Troy R. HAWKINS
	dataset_link	https://onlinelibrary.wiley.com/doi/abs/10.1111/j.15
	dataset_format	Excel spreadsheet
	project_report	DOI: 10.1111/j.1530-9290.2012.00532.x
	suggested_citation	DOI: 10.1111/j.1530-9290.2012.00532.x
	visible	1
Data submission, review, and conversion info	access_date	42281
	submission_date	2018-06-04
	submitting_user	Stefan Pauliuk
	dataset_conversion_info	Numbers converted from Excel to database via Pyt
	review_date	
	review_user	
	review_comment	
Reserve	reserve1	
	reserve2	
	reserve3	
	reserve4	
	reserve5	

dataset id and description

dataset aspects

dataset metadata

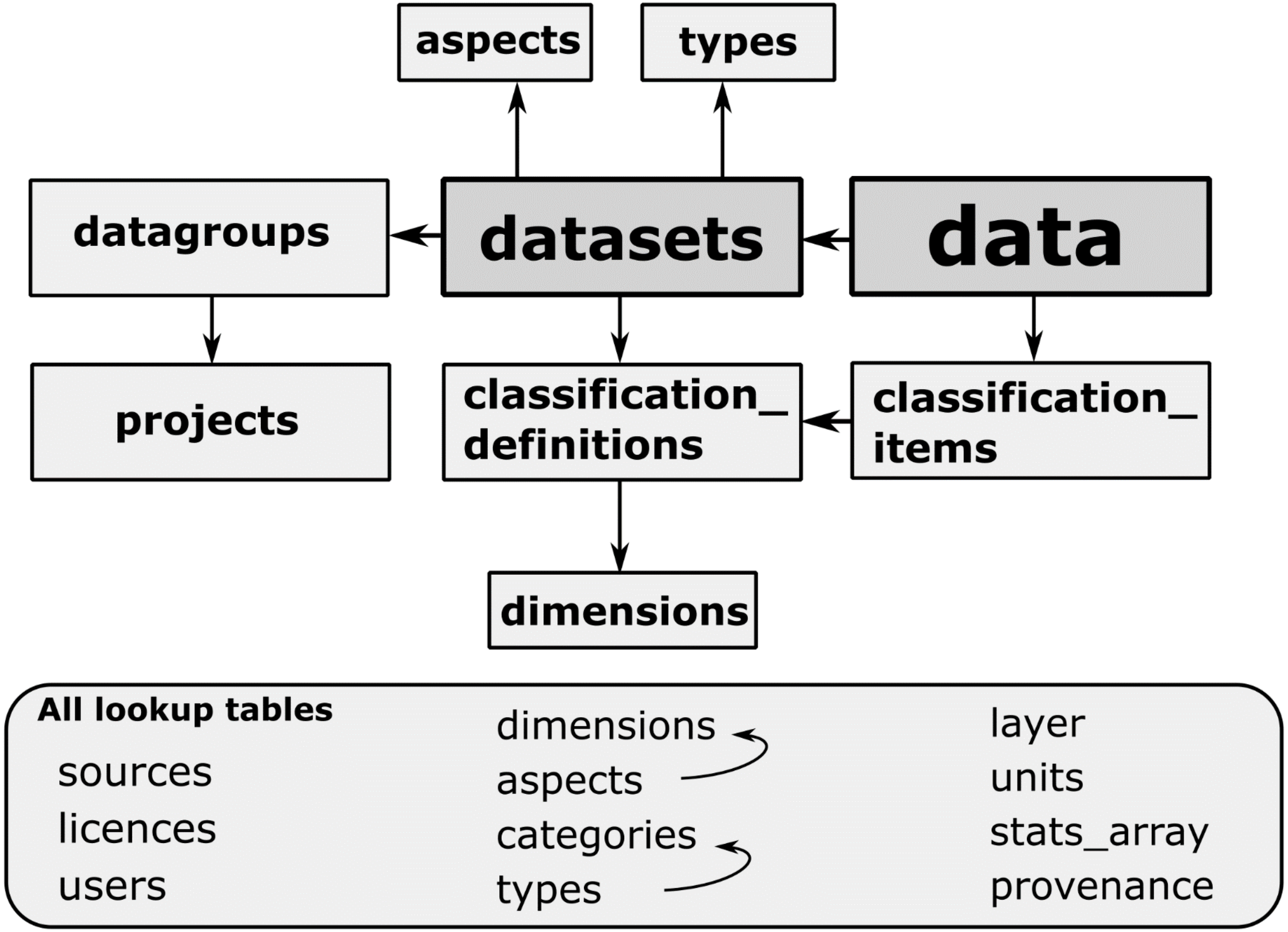
data table entry: List of tuples (table data are also possible)

material	commodity	region	age-cohort	value	unit
steel	passenger vehicle, internal combu	Global	2000-2010	882.0902775	kg
cast iron	passenger vehicle, internal combu	Global	2000-2010	115.8742446	kg
aluminium	passenger vehicle, internal combu	Global	2000-2010	71.2499941	kg
copper	passenger vehicle, internal combu	Global	2000-2010	23.87236603	kg
magnesium	passenger vehicle, internal combu	Global	2000-2010	0.240403956	kg
zinc	passenger vehicle, internal combu	Global	2000-2010	0.099790321	kg
lead	passenger vehicle, internal combu	Global	2000-2010	11.54087636	kg
gold	passenger vehicle, internal combu	Global	2000-2010	0	kg
silver	passenger vehicle, internal combu	Global	2000-2010	0	kg
neodymium	passenger vehicle, internal combu	Global	2000-2010	0	kg
platinum	passenger vehicle, internal combu	Global	2000-2010	0.0011	kg
paladium	passenger vehicle, internal combu	Global	2000-2010	0.0053	kg
rhodium	passenger vehicle, internal combu	Global	2000-2010	0.0006	kg
arsenic	passenger vehicle, internal combu	Global	2000-2010	0.004940321	kg
antimony	passenger vehicle, internal combu	Global	2000-2010	0.116920923	kg
plastic	passenger vehicle, internal combu	Global	2000-2010	186.3873028	kg

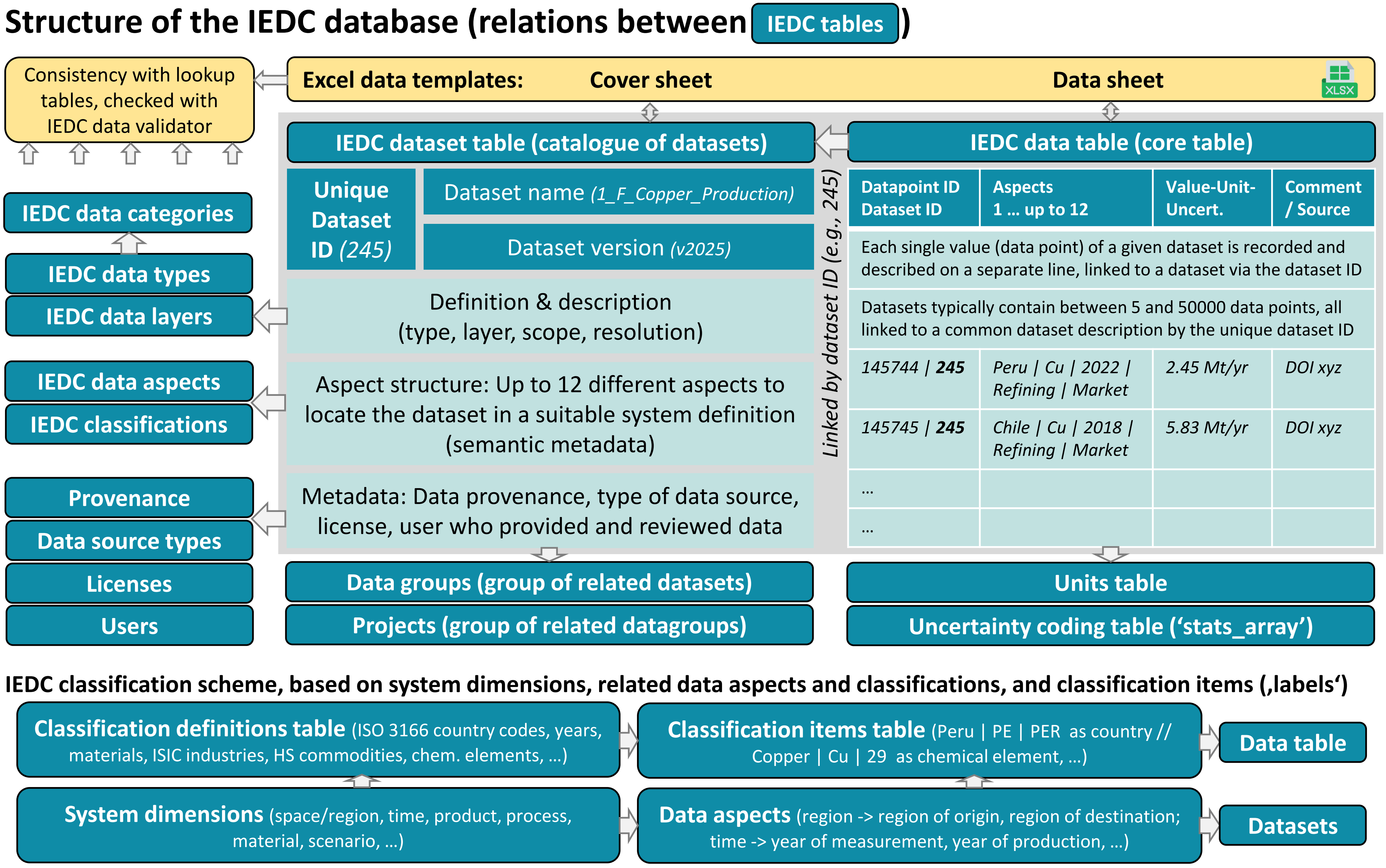
Aspects and classification items **Value** **Unit**

Source: „A General Data Model for Socioeconomic Metabolism and its Implementation in an Industrial Ecology Data Commons Prototype ", by Pauliuk et al. (2019), JIE, DOI: 10.1111/jiec.12890

Using the data model II



Source: „A General Data Model for Socioeconomic Metabolism and its Implementation in an Industrial Ecology Data Commons Prototype ", by Pauliuk et al. (2019), JIE, DOI: 10.1111/jiec.12890


Structure of the IEDC database (relations between IEDC tables)
Consistency with lookup tables, checked with IEDC data validator
Excel data templates: Cover sheet
Data sheet
XLSX
IEDC dataset table (catalogue of datasets)
Unique Dataset ID (245)
Dataset name (1_F_Copper_Production)
Dataset version (v2025)
Definition & description (type, layer, scope, resolution)
Aspect structure: Up to 12 different aspects to locate the dataset in a suitable system definition (semantic metadata)
Metadata: Data provenance, type of data source, license, user who provided and reviewed data
Linked by dataset ID (e.g., 245)
IEDC data table (core table)
Datapoint ID Dataset ID
Aspects 1 … up to 12
Value-Unit-Uncert.
Comment / Source
Each single value (data point) of a given dataset is recorded and described on a separate line, linked to a dataset via the dataset ID
Datasets typically contain between 5 and 50000 data points, all linked to a common dataset description by the unique dataset ID
145744 | 245
Peru | Cu | 2022 | Refining | Market
2.45 Mt/yr
DOI xyz
145745 | 245
Chile | Cu | 2018 | Refining | Market
5.83 Mt/yr
DOI xyz
…
…
IEDC data categories
IEDC data types
IEDC data layers
IEDC data aspects
IEDC classifications
Provenance
Data source types
Licenses
Users
Data groups (group of related datasets)
Projects (group of related datagroups)
Units table
Uncertainty coding table ('stats_array')
IEDC classification scheme, based on system dimensions, related data aspects and classifications, and classification items (‚labels')
Classification definitions table (ISO 3166 country codes, years, materials, ISIC industries, HS commodities, chem. elements, …)
Classification items table (Peru | PE | PER as country // Copper | Cu | 29 as chemical element, …)
Data table
System dimensions (space/region, time, product, process, material, scenario, …)
Data aspects (region -> region of origin, region of destination; time -> year of measurement, year of production, …)
Datasets

Schematic overview of the IEDC data categories (1-8) and data types (S, F, MC, LT, EI, UPI, …)

	Objects: goods, products, substances, commodities, waste	**Processes: industries, markets, end-use sectors, use-phase**
Extensive (at scale)	**Flows (1)** Flow (1_F_) | Process inventory (1_PI_) Births/Deaths (1_BD_) | Lorenz curve flows (1_LC_)	**Extensive process properties (5)** Process output capacity (5_CAP_) Process parameters (5_PP_)
	Stocks (2) Stock (2_S_) | In-use stock (2_IUS_) | Population (2_P_) Stock change (2_DS_) | Lorenz curve stocks (2_LC_)	
Intensive (per unit)	**Intensive object properties (3)** Product material composition (3_MC_) | Product lifetime (3_LT_) | Specific energy consumption (3_EI_) | Price (3_PR_) Shares (3_SHA_) | characterisation. factors (3_CF_)	**Intensive process properties (4)** Process yield factors (4_PY_) | Process environmental extensions per output (4_PE_) | Process operating costs per output (4_PC_) | unit process inventory (4_UPI_) | transfer coefficients (4_TC_)
General ratios	**General ratios (6)** Per capita stock (6_PCS_) | Per capita flows (6_PCF_) | Material substitution coefficient (6_MSC_) | miscellaneous intensive properties (6_MIP_) | criticality metrics (6_CR_) | impact indicators (6_IMP_)	
	(7) Classifications and correspondence tables: 7_CT	
	(8) Non-numeric data, Boolean data types: 8_FLG	

Data that don‘t fit into the IEDC data model: geospatial data, live-streamed data, event-based data like protocols, decision trees or flow diagrams, functional relationships, data with conditions such as thresholds, constraints across data points (summation constraints, Dirichlet distributions)

IEDC slide deck

History

IEDC history: The Yale Stocks and Flows database


Yale SCHOOL OF THE ENVIRONMENT | ABOUT US | CONTACT US | NEWS | SEARCH THIS SITE

Center for Industrial Ecology

Education | Research | Events | YSE Events | Media | People | Resources | Opportunities

HOME » RESEARCH » STOCKS AND FLOWS PROJECT (STAF)

Project Information

Principal Investigator(s):
Thomas Graedel

PI Website: https://environment.yale.edu/profile/graedel/

Funder: AT&T, NSF, Grainger Foundation, Intl Copper Association, Aluminum Association, International

Stocks and Flows Project (STAF)

Project Overview

Modern society is made possible by the use of metals, and metals have historically been supplied from virgin stocks (ore bodies, mineral deposits, and the like). Other reservoirs exist, however, a principal one being materials or products in use, stored, or discarded over the years by corporations and individuals. These reservoirs might become very important in the next few decades of rapid population growth and resource and energy use. There are also concerns about the use of energy in the extraction and processing of metals, and realization that the loss of resources by dissipation or landfilling can sometimes be problematic from an environmental standpoint, and concerns over the short and long-term “criticality” of metals.

https://cie.research.yale.edu/research/stocks-and-flows-project-staf

YALE-STAF: Systematic attempt to conduct and MFA for large parts of the periodic table

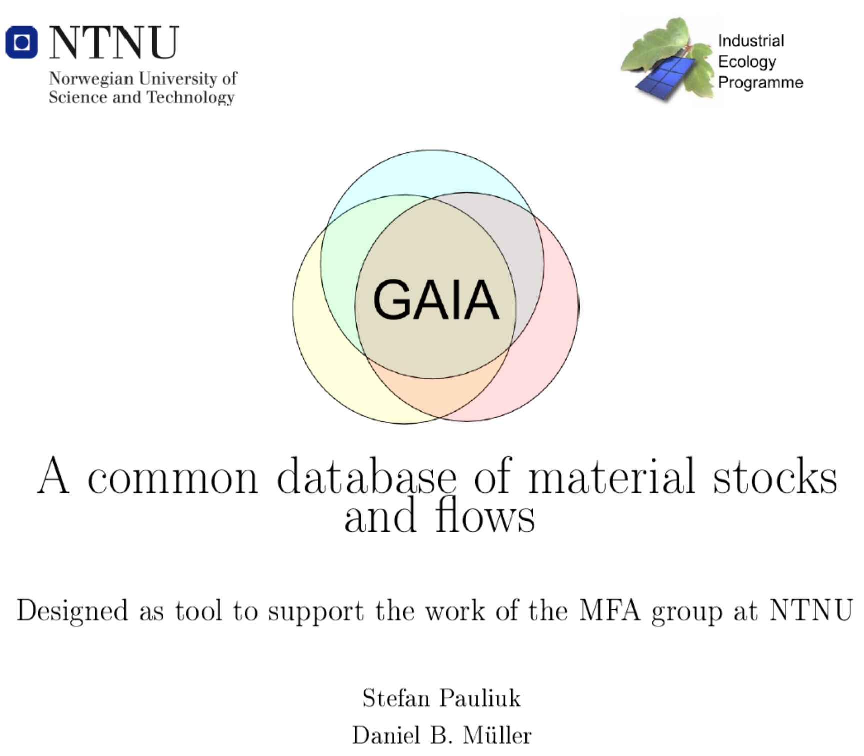
Need to have a systematic database arose early

Barbara Reck and Daniel B Müller worked on a first systematic attempt

Data from STAF project were published eventually by Myers et al. (2019)

Myers, Reck, and Graedel (2019): YSTAFDB, a unified database of material stocks and flows for sustainability science https://doi.org/10.1038/s41597-019-0085-7

IEDC history: The GAIA database at Indecol NTNU, Trondheim


NTNU
Norwegian University of Science and Technology

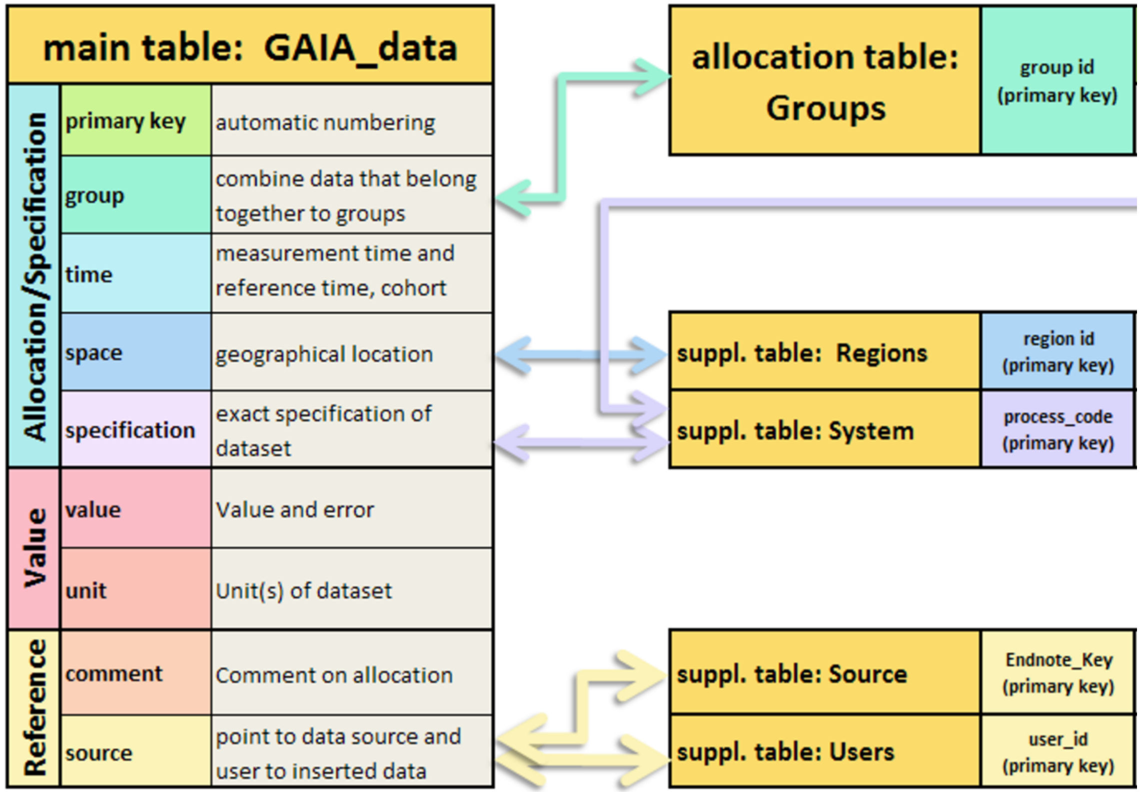
Industrial Ecology Programme

GAIA

A common database of material stocks and flows

Designed as tool to support the work of the MFA group at NTNU

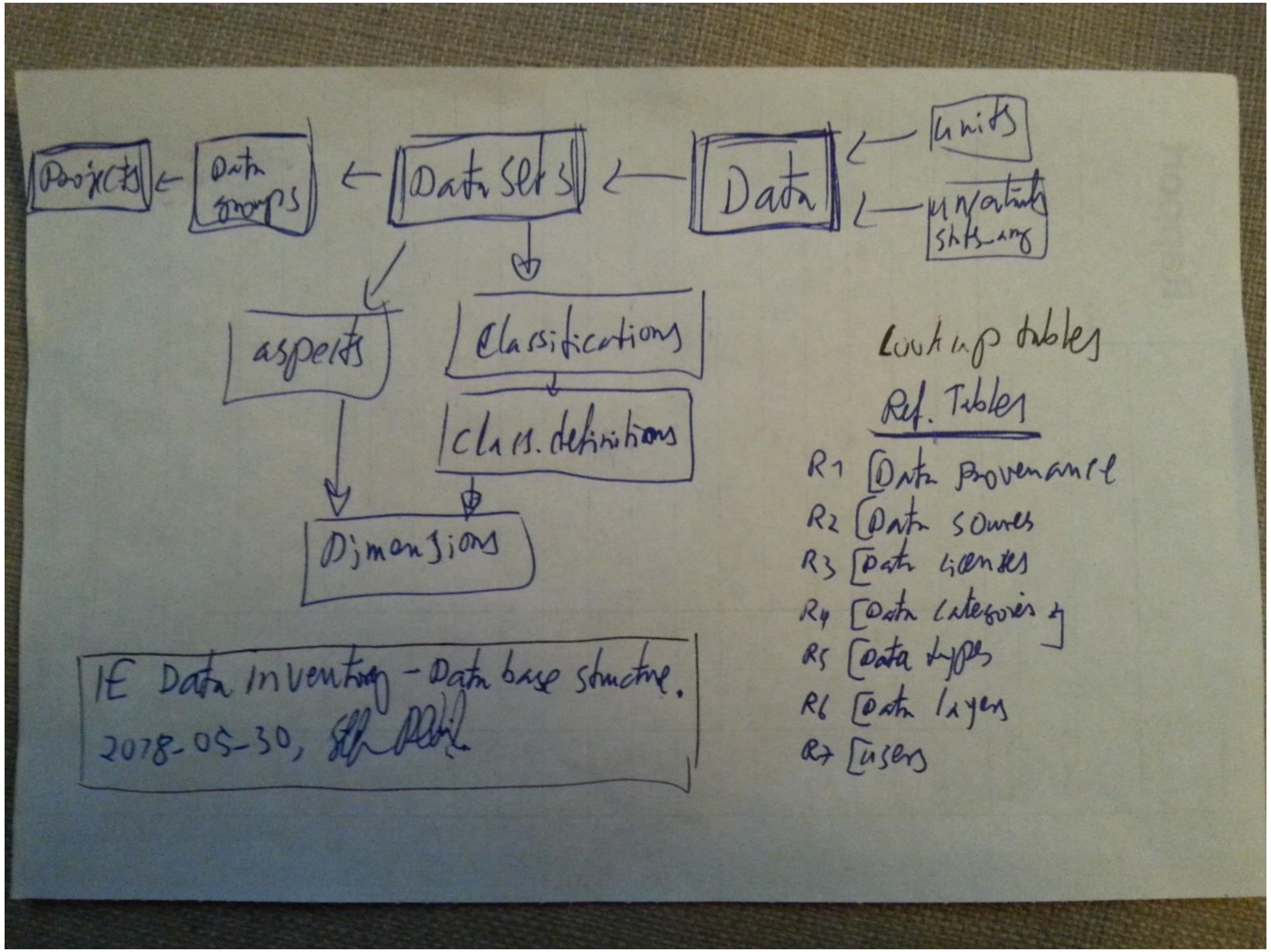
Stefan Pauliuk
Daniel B. Müller



- Daniel Müller brought the idea to Trondheim and continued pushing for a common stocks and flows database for MFA, where all data are formatted into a single large table (one row for each datapoint)
- Stefan Pauliuk implemented an SQL-based version of such a database as part of his ‚duty work‘ for the department: GAIA.
- GAIA contained a few common classifications (chem. elements, regions, processes), and had rather rigid aspects (5 time aspects, 3 space aspects, 4 process aspects, and 2 for product/material) and was used for flow and stock data only.
- This prototype fulfils some requirements of industrial ecology 2.0 (Davis et al. 2010): tailored data model, machine and human-readable data, open source tools, open data, shared vocabulary.

IEDC history: The Freiburg revival

- Stefan Pauliuk brought the idea to Freiburg and continued working on it as a ‚pet project‘.
- Further development of the data model to be generally applicable, link to established classifications
- Decision to not only build a catalogue of available data but to actually reformat and compile the data into a common database
- With Mahadi Hasan, we built the group‘s public dabatase as first implementation of the data model in 2018.
- Tested first with own research results, later expanded to broad coverage of IE data.



https://www.blog.industrialecology.uni-freiburg.de/index.php/2018/10/17/launching-the-prototype-for-an-industrial-ecology-data-inventory/

IEDC history: The general data model and the Freiburg prototype

JOURNAL OF INDUSTRIAL ECOLOGY

International Society for Industrial Ecology

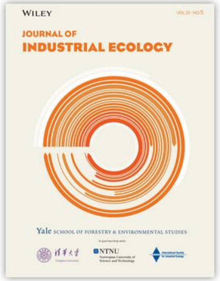
METHODS, TOOLS, AND SOFTWARE | Open Access

A general data model for socioeconomic metabolism and its implementation in an industrial ecology data commons prototype

Stefan Pauliuk, Niko Heeren, Mohammad Mahadi Hasan, Daniel B. Müller

First published: 06 May 2019 | https://doi.org/10.1111/jiec.12890 | Citations: 23

Editor Managing Review: Guillaume Majeau-Bettez

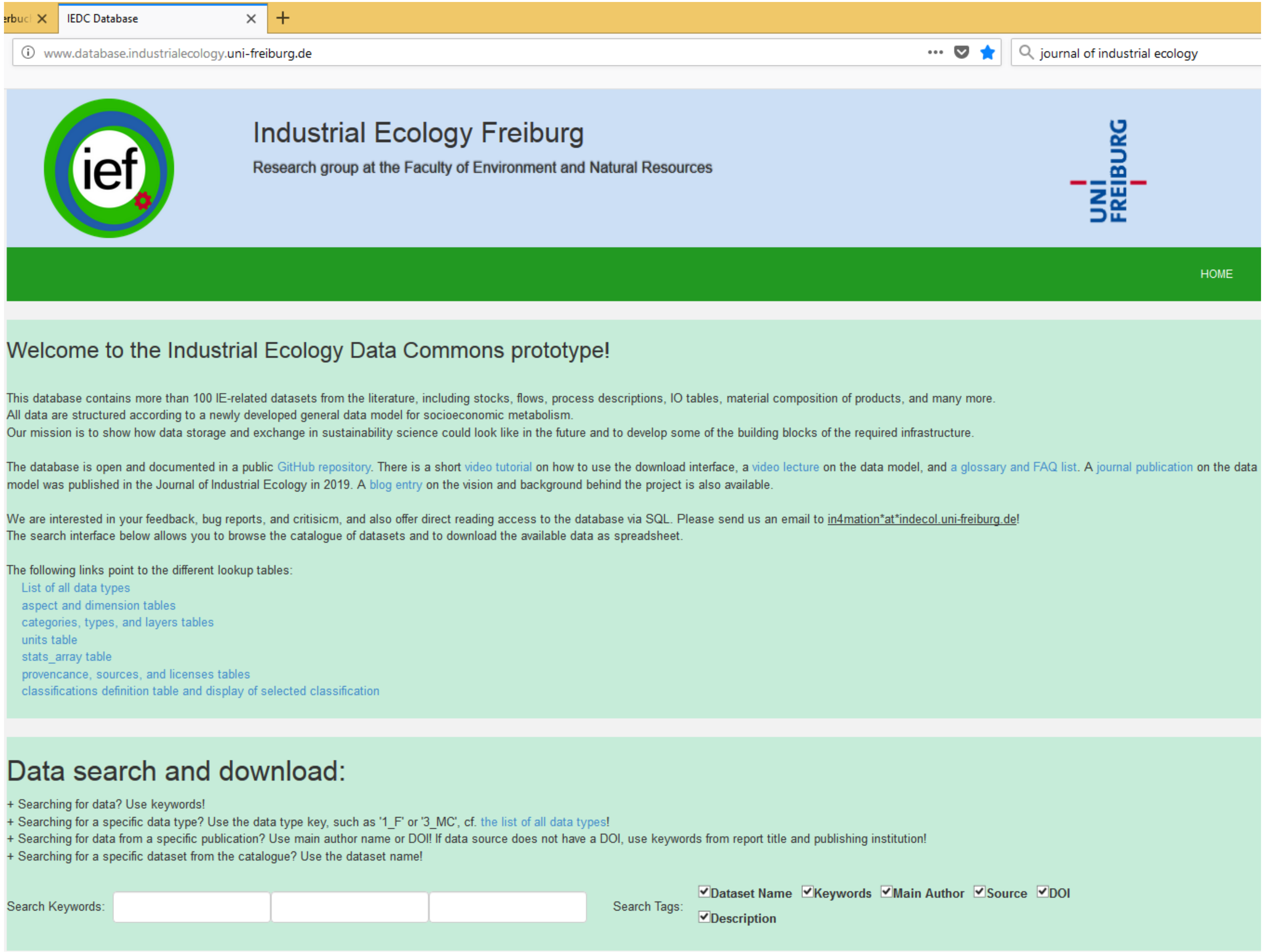


Volume 23, Issue 5
October 2019
Pages 1016-1027
This article also appears in:
JIE Articles in the IPCC 6th Assessment Report

- Published the data model and the structure of the database implementation as JIE article in 2019.
- Since 2019: regular updates (2x per year) to the iedc prototype, 30-40 new datasets each year.
- With Nildem Atasayar, we added new search and validation features.
- 2024: The *iedc prototype* became the IEDC.
- 2025: EU Horizon CIRCOMOD input data hub and IEDC Critical Mass Sprint flushed in heaps of new data!

https://doi.org/10.1111/jiec.12890

IEDC history: The first public interface, 2018-2024


erbuc IEDC Database

www.database.industrialecology.uni-freiburg.de

journal of industrial ecology

Industrial Ecology Freiburg

Research group at the Faculty of Environment and Natural Resources

UNI FREIBURG

HOME

Welcome to the Industrial Ecology Data Commons prototype!

This database contains more than 100 IE-related datasets from the literature, including stocks, flows, process descriptions, IO tables, material composition of products, and many more.
All data are structured according to a newly developed general data model for socioeconomic metabolism.
Our mission is to show how data storage and exchange in sustainability science could look like in the future and to develop some of the building blocks of the required infrastructure.

The database is open and documented in a public GitHub repository. There is a short video tutorial on how to use the download interface, a video lecture on the data model, and a glossary and FAQ list. A journal publication on the data model was published in the Journal of Industrial Ecology in 2019. A blog entry on the vision and background behind the project is also available.

We are interested in your feedback, bug reports, and critisicm, and also offer direct reading access to the database via SQL. Please send us an email to in4mation*at*indecol.uni-freiburg.de!
The search interface below allows you to browse the catalogue of datasets and to download the available data as spreadsheet.

The following links point to the different lookup tables:
- List of all data types
- aspect and dimension tables
- categories, types, and layers tables
- units table
- stats_array table
- provencance, sources, and licenses tables
- classifications definition table and display of selected classification

Data search and download:

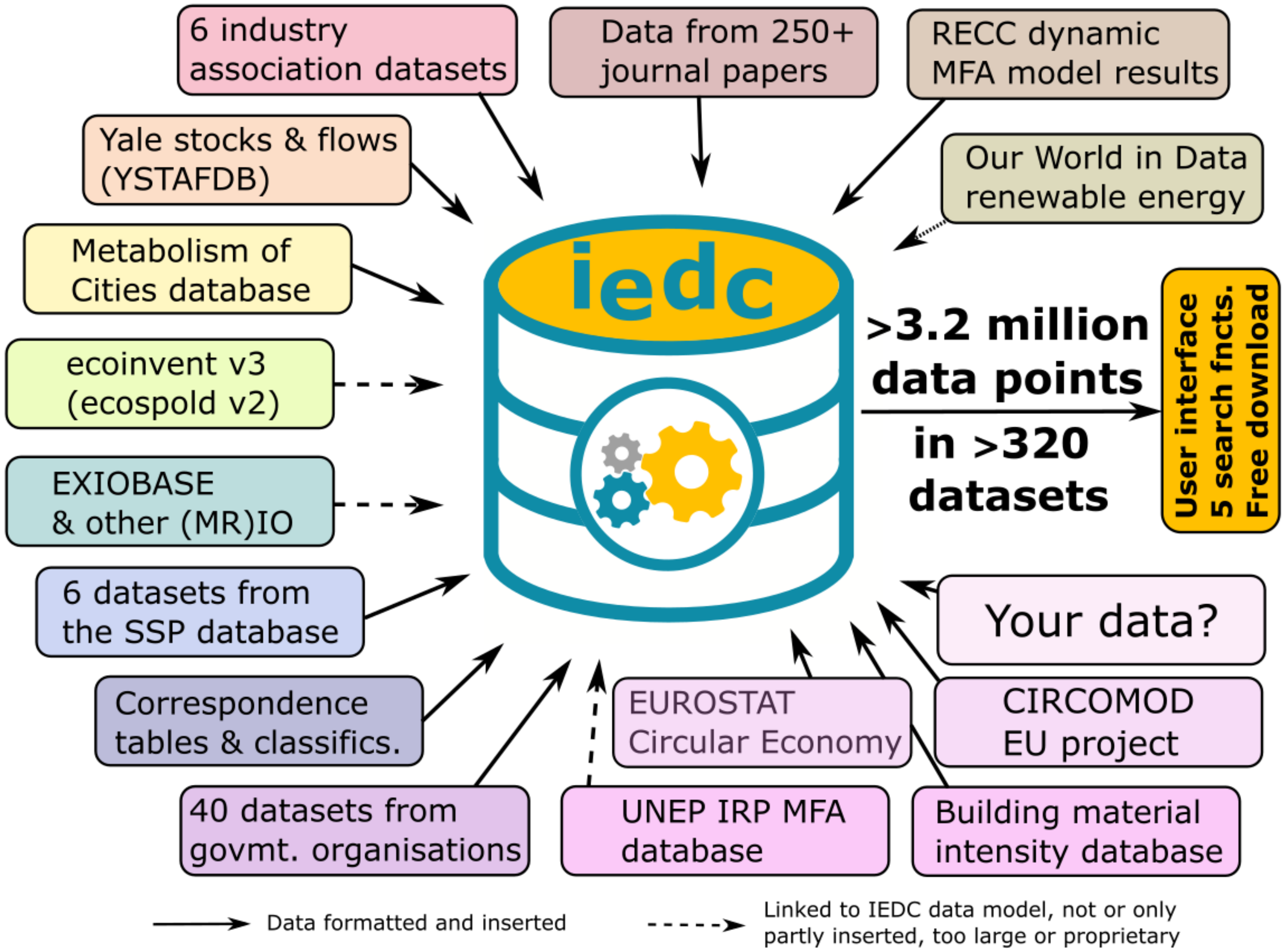
+ Searching for data? Use keywords!
+ Searching for a specific data type? Use the data type key, such as '1_F' or '3_MC', cf. the list of all data types!
+ Searching for data from a specific publication? Use main author name or DOI! If data source does not have a DOI, use keywords from report title and publishing institution!
+ Searching for a specific dataset from the catalogue? Use the dataset name!

Search Keywords: Search Tags: ☑Dataset Name ☑Keywords ☑Main Author ☑Source ☑DOI ☑Description

www.database.industrialecology.uni-freiburg.de

IEDC present: 2025 IEDC Critical Mass Sprint

- Collect, format, and upload a larger number (100+) of datasets on product material composition, energy intensity, and lifetimes
- Products include appliances, buildings, vehicles, infrastructure, industrial assets, and energy system technologies
- Critical mass sprint goal: By the end of 2025, the IEDC contains enough relevant data that searching the IEDC becomes a useful screening step / first iteration for MFA and LCA researchers and consultants



iedc **Critical mass sprint 2025**

https://www.database.industrialecology.uni-freiburg.de/
https://www.blog.industrialecology.uni-freiburg.de/index.php/2025/03/03/2025-iedc-critical-mass-sprint/

IEDC future: 2026 and beyond

- Offer seminars, documentation, and other training material to promote the use of the IEDC (both upload and download), the general data model behind it, the data templates, and the IEDC classifications

- Continue adding data to establish the IEDC as a platform for a broad spectrum of socio-metabolic and industrial systems data. Focus on the following data types:
  - Material composition (3_MC), specific energy consumption (3_EI), and lifetime (3_LT) of products, buildings, vehicles, and infrastructure, and related process descriptions (4_PY, 4_PE, 4_UPI)
  - Sankey diagrams and other MFA stock and flow data (1_F, 2_S, 2_IUS)
  - Inequality-related data (Lorenz curves for flows, stocks, and normalized: 1_LCF, 2_LCS, 3_LC)

- Up next: explore the use of AI-based tools to gather and format data
  - Need to ensure accuracy and traceability for the data gathered

https://www.database.industrialecology.uni-freiburg.de/

IEDC logo and colors

For the buttons and tables, please adhere to the IEDC colour scheme:
Orange: (255,192,0), (255,229,147)
Blue-green: (15,140,167), (193,225,224)

https://www.database.industrialecology.uni-freiburg.de/