Object-Centric Process Mining for Semi-Automated and Multi-Perspective Sustainability Analyses

Nina Graves*†, Andreas Fritsch‡, Raimund Hensen*§, István Koren*, and Wil M. P. van der Aalst*
*RWTH Aachen University, Chair of Process and Data Science, Aachen, Germany
‡Karlsruhe Institute of Technology, Institute AIFB, Karlsruhe, Germany
Email: †graves@pads.rwth-aachen.de, ‡andreas.fritsch@kit.edu, §raimund.hensen@rwth-aachen.de

***Abstract*—Current approaches for sustainability analysis face significant methodological challenges in comprehensively assessing impacts across business processes. They are constrained by time-consuming data collection and allocation, inflexible, manually created models, and difficulties in relating impacts to root causes. This paper considers Object-Centric Process Mining (OCPM) as enabler for more dynamic and comprehensive sustainability analyses mitigating these challenges. We explore the integration of sustainability data into the current standard for Object-Centric Event Logs (OCELs) used by OCPM techniques and introduce a prototypical tool to assess impacts based on an OCEL and allocate these impacts to events and objects. We demonstrate how OCPM can support the specification of inventory data, impact assessment, and impact allocation for a semi-automated, multi-perspective sustainability analysis. This exploratory work provides a concrete illustration for sustainability assessment using OCPM, thereby, laying the foundation for further research towards a data and process driven technique to identify and measure company's negative sustainability impacts more effectively.**

***Index Terms*—process mining, sustainability, object-centric process mining, life cycle assessment, data format**

I. Introduction

Facing the intensifying global challenges, such as the climate crisis, requires companies to put increasing emphasis on the reduction of their operation's negative impacts on people and the environment [11]. Sustainability analysis methods and guidelines such as Life Cycle Assessment (LCA) [26] and the Greenhouse Gas Protocol [39] guide companies in identifying and measuring such negative impacts. Nevertheless, company's face significant difficulties in the application of these time-consuming and costly methods and guidelines [2], [7]: Most sustainability analyses are based on manually created models and data in different degrees of aggregation collected from multiple sources that has to be combined. Hence, the models describe a static snapshot of a complex dynamic system. The resulting data and model basis is inflexible and the required maintenance for ongoing analysis becomes increasingly difficult. Finally, there is a gap between reporting practises and the effective use of sustainability data to manage and reduce the negative impacts of companies' operations [11].

Process Mining (PM) is a relatively new field that emerged from the realisation that the analysis of business processes based on manually created models are unreliable and inaccurate [36]. Instead, PM techniques use data logged during the execution of business operations to model, analyse, monitor and improve business processes [36]. PM has proven highly beneficial in industrial applications, leading to significant reductions in costs and increases in efficiency and quality [34]. This resulted in two trends: 1) More and more companies start managing their event data to use PM for data-driven process management [16], and 2) Event data has evolved from a one-time export of a single IT system's event log describing one process to a data structure that combines data from multiple systems as a single source of truth [3]. Object-Centric Process Mining (OCPM) is a recent development in PM that leverages such multi-perspective event data for a holistic analysis of a landscape with several, interconnected business processes [37].

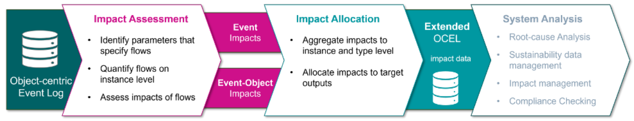


Figure 1: Before OCPM techniques can support the analysis of impact drivers, the OCEL has to support impact assessment and allocation.

While PM's potential for sustainability analysis has been acknowledged in multiple works [6], [15], [18], [34] the concrete implementation remains underexplored. In this paper, we explore impact assessment and allocation based on an Object-centric Event Log (OCEL) as depicted in Figure 1. We describe a general OCEL-based approach to sustainability analysis, consisting of *Impact Assessment*, *Impact Allocation* and *System Analysis* using an illustrative example. Furthermore, we provide a web application as a proof of concept for creating a sustainability-enriched OCEL and offer an elaborate discussion of the innovative analysis capabilities, the challenges and limitations.

We offer an overview of existing related approaches and how our work differs from and extends them in Section III. Section II, introduces our running example and provides background information on LCA as a central sustainability analysis method and OCPM. In Section IV, we explore the combination of OCPM and LCA and consider how sustainability data can be integrated in OCEL. We show the technical feasibility of our approach using a prototypical implementation, presented in Section V. Contributions and limitations of our work are discussed in Section VI. We conclude with an outlook to future work in Section VII.

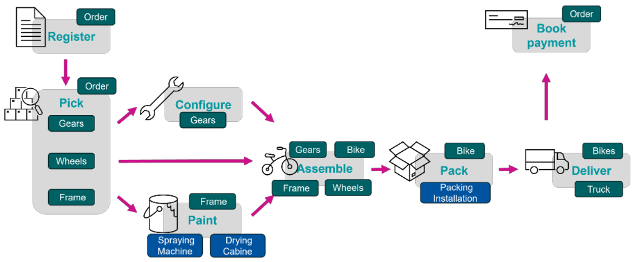


Figure 2: Example process.

II. Preliminaries

We introduce an illustrative example as well as background information and notations for the LCA standard, which we use as a reference for sustainability analysis, and OCPM.

A. Running Example

Figure 2 represents the running example process for custom-made bike production. The bigger, grey boxes depict activities and the smaller boxes attached to them represent participants involved in the activities; green boxes represent handling units (entities we are working *on*), and blue boxes are process resources (entities we are working *with*). The process begins with the *registration* of an order which specifies what parts need to be *picked* from the warehouse. Next, the gears are *configured*, and the frame is *painted*, before the bike is *assembled*. After assembly, the bike is *packed* and *delivered*, and the booking of the *payment* takes place.

B. Life Cycle Assessment

Analysing the social and environmental impacts of corporate activities is a complex challenge. A large number of different sustainability aspects (different environmental and social issues) must be included in an analysis. Furthermore, sustainability impacts of an activity may also be located outside the company. An important concept in this context is life cycle thinking [24]. Life cycle thinking involves a systemic perspective, which considers the relationships and interactions between a product or company and its environment as well as other stakeholders. The goal is to identify opportunities for improvement within the system as a whole [23]. The application of the life cycle approach is considered essential for the study of corporate sustainability [10], [32]. Life Cycle Assessment (LCA) is a standardized method for conducting such an investigation [26]. The LCA method was initially developed to analyse the environmental impact of products [20]. Recent developments of the method broaden the perspective to analyse the environmental impact of an entire company [27]. Furthermore, approaches now exist at both the product and corporate level that also take into account the social dimension of sustainability [20].

According to [26], the LCA method defines four phases of a sustainability analysis. Phase 1 of an LCA is the definition of **goal and scope**. This definition must describe the considered impact categories (e.g., climate change, mineral resource scarcity,..), the total system's reference flow and the considered system boundaries (e.g., cradle-to-gate, gate-to-gate,..) as well as basic requirements and assumptions. It usually includes the manual creation of a system model. The model describes all of the operations performed in the context of the reference unit and could look similar to the one displayed in Figure 2. In our example, we assume that the company is interested in analysing climate change impacts from cradle to gate (i.e. including the upstream supply chain).

The second phase is the **inventory analysis**, where data is collected and calculations are made to quantify sustainability-relevant *flows* (inputs and outputs). Using the system model, each operation's flows relevant to the selected impact categories are determined and quantified.

Example: The operation paint frame has three flows: It uses a frame (flow 1) as input and requires energy to power the machines involved (flow 2). Furthermore, the output is a single painted frame (flow 3) which must be associated with an impact to be able to determine the impact of the final bike. For the sake of simplicity, we neglect other potentially relevant flows.

All flows need to be quantified relative to the reference flow. Depending on available company data and context information, this may require additional derivations.

Example (cont.): We only know that the machines use 615 kWh per month on average. Using the average number of frames painted per month (6), we know that the average frame's energy consumption lies at $615\ [\frac{kWh}{month}]\ /\ 6\ [\frac{frame}{month}] = 102.5\ [\frac{kWh}{frame}]$.

Based on the identified flows and the collected data, the impact of each flow is determined in the **impact assessment** phase. This phase combines two major tasks: the determination of an impact and the allocation of the impact to the analysis' reference flow. Impacts can typically be derived as secondary data from LCA databases.

Example (cont.) Using an LCA database and the inventory data, we determined each frame's impact, so we know the impact of flow 1 is $29.6\ [\frac{kg\ co_2e}{frame}]$. *Our factory lies in Germany, so we use the impact factor for 1 kWh of energy with the German energy mix from a database to determine the impact caused by the used energy:* $102.5\ [kWh/frame] \cdot 1.304\ [\frac{kg\ co_2e}{kWh}] = 133.7\ [\frac{kg\ co_2e}{frame}]$. *As flow 3 refers to a single frame and we only need one frame per bike (our reference flow), we do not need to allocate the total impact to multiple flows or multiple frames, i.e., the impact per painted frame is* $29.6\ [\frac{kg\ co_2e}{frame}] + 133.7\ [\frac{kg\ co_2e}{frame}] = 163.26\ [\frac{kg\ co_2e}{frame}]$.

The final phase is the **interpretation** of the results, where conclusions are drawn and changes to the process proposed. Phases are not necessarily conducted strictly sequentially. LCA is also an iterative approach, so that, for example, the most significant problem areas are identified in one iteration. For these problem areas, the data quality requirements for an in-depth analysis are then tightened.

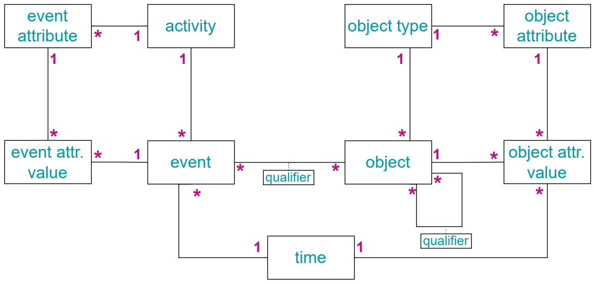


Figure 3: Meta Model of the Object-Centric Event Log Standard OCEL 2.0 [3].

C. (Object-centric) Process Mining

PM is a data-driven bottom-up approach to business process analysis. The central point of reference is an *event log* that is extracted from IT systems supporting business operations which documents every *event* that occurs. An event is a specific execution of an *activity* and has a timestamp informing on the date and time the event was executed. Different activities can have different properties that can differ from execution to execution, e.g., the brand of paint used in a painting activity or the travelled distance of a delivery; we refer to them as *attributes* and their *attribute values*. An event can be associated with one or more *objects* that are involved in the event. For example, we register an order called "o1" in one event of activity "Register" and order "o2" in the next. Every object is of a specific *object type* that describes the nature of an object and groups objects that have the same role in a process, e.g., both "o1" and "o2" are of object type order. As with events, two objects can be of the same type but differ in their attribute values, e.g., bike1 is a racing bike while bike2 is a gravel bike. Such object attribute values may also change over time. Furthermore, two objects can share a specific relationship, for example, bike1 can be tied to order o1. Figure 3 depicts the meta model of the current OCEL 2.0 standard which depicts this understanding of a system.

Table I shows a simplified OCEL that could be extracted from the IT systems of a manufacturing compan and used for PM. The OCEL is represented in three tables: On the left, events and objects are given with identifiers, activity resp. object type, and selected further attributes. On the right, the E2O relations are listed showing what objects are involved in which events, i.e., event ev2 refers to order o1, frame1, wheels1 and gears1.

III. Related Work

Our proposal is related to approaches in different disciplines. On the one hand, it is related to efforts to automate data collection and calculation in *LCA*. On the other hand, our approach builds upon efforts within the *process science* discipline to incorporate sustainability aspects.

Related approaches to automate data collection and calculation in LCA aim to alleviate the time-consuming and costly nature of conventional (manual) sustainability analyses. Recent proposals to automate LCA (e.g. [19], [22], [35]) share several characteristics with our work. In particular, they are data-driven, strive to streamline data collection and calculation, as well as providing more transparency and traceability for analyses.

In contrast to these approaches, our work draws from process science—a field that encompasses areas like Business Process Management (BPM) and PM, focusing on analysing and improving business processes to enhance organizational performance [38]. While, traditionally, process science has been concerned with conventional performance measures such as time, cost and efficiency, more recently, researchers have started to incorporate sustainability aspects [8], [9], [15]. A major argument for developing sustainability analysis approaches from a process science perspective in addition to LCA-based approaches is its focus on management and improvement of processes [15]. Established BPM and PM tools like SAP Signavio[1], Camunda[2] or Celonis[3] are management tools, intended to e.g. provide decision support for managers or automate the execution of processes. The successful integra-

Table I: A simplified OCEL from the example process.

eID	activity	time	dist. [km]
ev1	Register	12-10 12:21	⊥
ev2	Pick Items	12-10 14:38	⊥
ev3	Configure	12-10 15:46	⊥
ev4	Paint Frame	12-10 16:06	⊥
ev5	Register	15-10 10:15	⊥
ev6	Pick Items	15-10 13:42	⊥
ev7	Paint Frame	15-10 15:52	⊥
ev8	Assemble	16-10 15:30	⊥
ev9	Configure	17-10 12:32	⊥
ev10	Pack Bike	17-10 15:28	⊥
ev11	Assemble	18-10 09:46	⊥
ev12	Pack Bike	18-10 16:14	⊥
ev13	Deliver	19-10 10:42	256.7
ev14	Book Pay	21-10 12:05	⊥
ev15	Book Pay	24-10 17:04	⊥

oID	object type	bike type	weight [kg]
o1	Order	⊥	⊥
f1	Frame	⊥	⊥
w1	Wheels	⊥	⊥
g1	Gears	⊥	⊥
o2	Order	⊥	⊥
f2	Frame	⊥	⊥
w2	Wheels	⊥	⊥
g2	Gears	⊥	⊥
bk1	Bike	Race	11.4
bk2	Bike	Gravel	19.6
t1	Truck	⊥	⊥
sm1	Spray. In.	⊥	⊥
sm2	Spray. In.	⊥	⊥
dc1	Dry. Cab.	⊥	⊥
pk1	Pack. In.	⊥	⊥

eID	oID
ev1	o1
ev2	o1
ev2	f1
ev2	w1
ev2	g1
ev3	g1
ev4	f1
ev4	sm1
ev4	dc1
ev5	o2
ev6	o2
ev6	f2
ev6	w2
ev6	g2
ev7	f2
ev7	sm2
ev7	dc1
ev8	f1
ev8	w1
ev8	g1
ev8	bk1
ev9	g2
ev10	bk1
ev10	pk1
ev11	f2
ev11	g2
ev11	bk2
ev12	bk2
ev12	pk1
ev13	bk1
ev13	bk2
ev13	t1
ev14	o1
ev15	o2

[1]https://www.signavio.com/de/
[2]https://camunda.com/
[3]https://www.celonis.com

tion of sustainability aspects into process science approaches therefore provides a valuable complement to existing LCA-based approaches. Particularly, such sustainability-enhanced process science approaches may facilitate the development of ICT tools that help companies transition from merely reporting sustainability performance to actively managing and mitigating negative sustainability impacts [11].

Note that the distinction between LCA and process science approaches is based on both being disciplines that have, historically, developed independently from each other [5], [21]. Conceptually, there exist overlaps between both disciplines and we do advocate an integration of both perspectives [14]. Compared to LCA approaches, however, process science approaches described in literature are less comprehensive and elaborate in their conceptualization of sustainability impacts [14], [15]. Our approach addresses this issue by explicitly building upon LCA concepts. Furthermore, existing process science approaches have mainly focused on modelling concepts, without support for automated and data-driven analyses [15]. More closely related to our work are PM approaches which we describe in the following in more detail.

The framework presented in [33] shows how the phases of the LCA method (goal and scope definition, inventory analysis, etc.) can be mapped to generic steps necessary for process mining (planning, data extraction, data processing, ...). The authors investigate the capability of different PM algorithms to discover process models from manufacturing process logs and argue that these logs can be enriched with flow data (energy, material). Our work extends this by proposing a concrete structure for such enriched event logs and exploring analysis capability based on them.

Another framework proposed in [25] suggests the use of process mining to measure sustainability performance based on the GRI[4] sustainability *reporting* standard. Different to their approach, our proposal builds upon LCA as a comprehensive and systematic sustainability *analysis* method. This, however, is not incompatible with the approach of [25], since the LCA method can be used to conduct analyses that inform a report according to the GRI standard.

In an empirical study [6] the potential for using PM for carbon accounting is explored based on expert interviews. The authors conclude that process mining can increase transparency for carbon accounting and allow for more granular analyses.

The SOPA-framework [28] combines process mining techniques with life cycle assessment, thereby also posing a data and process driven approach. It touches all phases of the BPM-lifecycle, including process redesign, allowing for *what if* considerations to determine what changes in the process and its element have to be made to reduce the overall impact. The authors use activity-based costing, thereby, considering impacts as costs caused by the execution of an activity. In contrast to the approach in this work, SOPA is based on traditional process mining which is equivalent to using an OCEL that only has a single object type, and static object attributes [37].

[4]https://www.globalreporting.org

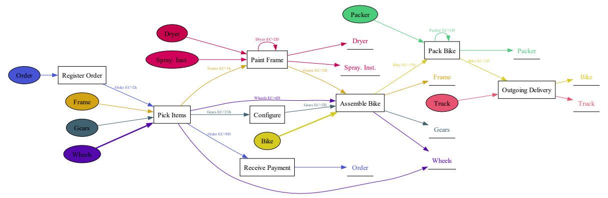


Figure 4: Process model (object-centric directly-follows graph) automatically derived from the OCEL using PM4Py [4]. The arc weights show the average time between events of the corresponding activities.

Overall, the idea of applying PM for sustainability analysis is relatively young, with some initial approaches to be found in the literature. Particularly, the object-centric paradigm has so far not been applied to sustainability analysis in existing PM approaches. Furthermore,the approach presented in the following is both process and data driven—this distinguishes it from recent efforts in the field of Green BPM. While Green BPM considers sustainability in the context of business processes [15], they are usually not data driven and lack concreteness [18]. The approach we present extends existing work by proposing a conceptual integration of LCA and PM, describing a concrete example of how to structure and work with OCEL data for sustainability analyses and providing a prototypical implementation of our approach in a software tool. We hereby build upon and elaborate on the general framework outlined in [18] and concepts proposed in [14].

IV. Sustainability Analysis using OCPM

In this section, we outline a general procedure for conducting sustainability assessment based on an OCEL. We consider the integration of sustainability-relevant data into OCEL 2.0, identify challenges and highlight differences to common sustainability assessment methods. For easier comparability, we relate this section to the phases of the LCA.

Goal and Scope: We assume that data in the event log is captured according to a predefined goal and scope. The goal is to use an OCEL to determine impacts on the instance level (events and objects) and to allocate these impacts to the system's relevant outflow. We can use the constantly increasing landscape of OCPM tools [30] to gain a general understanding of the process and the data, e.g., *OCPM*[5] or PM4Py [4] to automatically discover process models and Ocelot[6] to help overview the data. Figure 4 shows a simple process model automatically detected from the example OCEL. Note, that every object type is merely visualised with a start and a sink node; this is not representative to whether the object is an input or an output.

[5]https://www.ocpm.info
[6]https://ocelot.pm/

Table II: Descriptions of the input flows: frames, wheels, gears and paint.

(a) Impact information in object attributes.

oID	obj. type	AlMg1SiCu [g]	recyc.	kg co2e/kg	co2e [kg]
f1	Frame	1500	no	16	
w1	Wheels				23
g1	Gears				18
f2	Frame	2200	yes	7	
w2	Wheels				27
g2	Gears				15

(b) Event attributes for the activity *paint frame*.

eID	spraying dur. [min]	drying dur. [h]	paint	used paint [l]
ev-4	74.00	12.00	brand1	0.40
ev-8	55	7.5	brand2	0.23

Table III: Additional event and object attributes.

(a) Event attributes related to energy consumption.

eID	activity	energy consumption [kWh]
ev1	Register Order	0.000566
ev5	Register Order	0.000556
ev8	Assemble Bike	0.8
ev11	Assemble Bike	0.6
ev14	Book Payment	0.000558
ev15	Book Payment	0.000558

(b) Object attributes related to energy consumption.

oID	object types	energy cons. [kWh/h]	energy cons. [kWh]
sm1	Spray. Inst.	55	⊥
sm2	Spray. Inst.	70	⊥
dc1	Dry. Cabine	5	⊥
pk1	Pack. Inst.	⊥	0.21

Inventory Analysis: Information about flows can be captured in event and object attributes and their values. We see different versions of describing flows by considering the inventory data for the used paint and the frame in Table II; flow-information referring to the frame can be found in the object attributes described in Table IIa, whereas the flow of the paint is described in the event attributes Table IIb. The combination of the attributes *AlMg1SiCu [g]* and *recyc.* and their values specify and quantify the flow of the frame. The flow is specified by the type of material provided in the attribute name, whether or not the material is recycled or not (in attribute name and value) and the mass as reference unit (also in the attribute name). The flows are quantified in the attribute values.

In general, flows may be associated with 1) events and specified in event attributes (e.g. the energy consumption of events specified for Table IIIa), 2) objects and specified with object attributes (e.g. the energy consumption of the packing installation for the time period the OCEL captures in Table IIIb), or 3) both events and objects, specified by the combination of the attributes for the corresponding E2O relations. For example, the hourly energy consumption of the spraying installation and the drying cabine is specified in the object attributes (Table IIIb), and the duration is an event attribute in Table IIb. We see that inventory data can be both stored in the OCEL and used as basis to quantify flows.

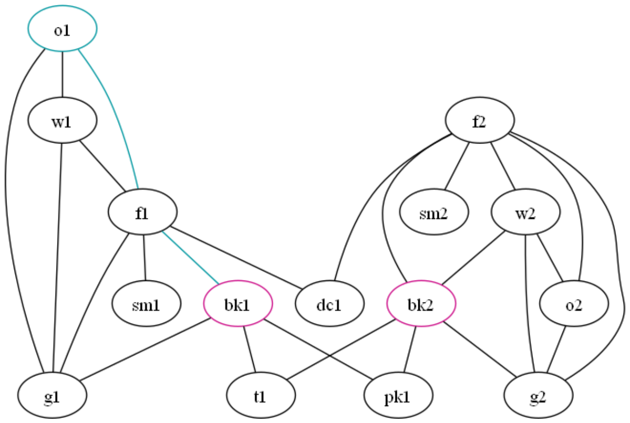


Figure 5: Object graph for the example process. The target objects, bikes, are pink. A shortest path from any object of event *ev1* to a target object is highlighted.

Impact Assessment: The impacts and impact factors of an event or object flow can be stored in the OCEL as an attribute. We see this in Table IIa, where the impact of the wheels and the gears and the impact factor for each frame is provided.

If the inventory data is provided in the log, we can determine impacts for events, objects and event-object combinations. For flows where the impact factors are not provided in an OCEL, the impact factor can be taken from an impact database based on the specification of the flow in the log. We do this by quantifying the flows based on the attributes and multiplying the result with the corresponding impact factor. For example, given an impact factor for consumed energy per kWh in Germany, we can determine the impact of each of the assembly operations, as the energy consumption is provided as event attribute. Table IV provides an overview of how the individual flows, impact factors and impacts can be derived based on the data in the OCEL.

With the current OCEL 2.0 standard, inventory data, impact factors and impacts tied to a single event and object can be stored and can be determined using an OCEL. However, flows and impacts tied to the combination of event and objects (e.g. impacts associated with using the spraying installation object in a painting event) cannot be stored in the OCEL directly as there are no E2O attributes. Flows and impacts are only derivable from the event and object attribute values.

In addition to determining all impacts for events and objects, sustainability assessment requires allocating impacts to selected outgoing flows, such as the bikes in our example. In general, these target outgoing flows could be associated with either an event or an object. OCPM can support the allocation of impacts by using the relationships among individual objects. Figure 5 shows an object graph that can be automatically derived from the OCEL. Every node is an object in the OCEL; every arc represents a shared event, i.e., an event to which both objects have an E2O relation. For example,

Table IV: Overview of inventory and impact data captured in and computed from the example OCEL. Highlighted attributes and values indicate that they influence the selection of the impact factor, and pink attributes specify an impact directly.

impact type description			reference entity					inventory data					impact assessment impact factor			impact assessment	
flow type	**impact name**	**scope**	**type**	**activity**	**object type**	**event**	**object**	flow description **relevant attributes**	units **corr.**	**origin**	flow **quantif.**	**unit**	**origin**	**value**	**unit**	**origin**	**value**
material	Frame	3	object		Frame			AlMg1SiCu [g] kg co2e / kg recyc.	$\frac{kg}{g}$	attr. value.	1.5 2.2	kg	attr. value	16 7	$\frac{kg\ co2e}{kg}$	computed	24 15
material	Wheels	3	object		Wheels		w1 w1	co2e / kg		none	⊥		none	⊥		attr. value	23 27
material	Gears	3	object		Gears		g1 g2	co2e / kg		none	⊥		none	⊥		attr. value	18 15
material	Paint	3	event	Paint		ev4 ev7		used paint[l] paint		attr. value.	0.4 0.23	l	impact DB	3.14 2.27	$\frac{kg\ co2e}{l}$	computed	1.256 0.522
energy	Registration	2	event	Register		ev1 ev5		energy consumption [kWh]		attr. value.	0.000566 0.000556	kWh	impact DB	1.304	$\frac{kg\ co2e}{kWh}$	computed	0.001 0.001
energy	Payment	2	event	Book Pay		ev14 ev15		energy consumption [kWh]		attr. value.	0.6 0.8	kWh	impact DB	1.304	$\frac{kg\ co2e}{kWh}$	computed	0.782 1.043
energy	Assembly	2	event	Assemble		ev8 ev11		energy consumption [kWh]		attr. value.	0.000558 0.000558	kWh	impact DB	1.304	$\frac{kg\ co2e}{kWh}$	computed	0.001 0.001
energy	Packing Instt.	2	object		Packer		pk1	energy cons. [kWh]		attr. value	0.21	kWh	impact DB	1.304	kg co2e / kWh	computed	0.274
energy	Spraying	2	E2O	Paint	Spray. Inst.	ev4 ev7	sm1 sm2	spraying dur. [min] energy cons. [kWh/h]	$\frac{h}{min}$	computed	67.83333 64.16667	kWh	impact DB	1.304	$\frac{kg\ co2e}{kWh}$	computed	88.455 83.673
energy	Drying	2	E2O	Paint	Dry. Cabine	ev4 ev7	dc1	drying dur. [h] energy cons. [kWh/h]		computed	60 37.5	kWh	impact DB	1.304	$\frac{kg\ co2e}{kWh}$	computed	78.240 48.900
pollution	Transport	1	E2O	Deliver	Bike	ev 13	bk1 bk2	dist. [km] weight [kg]	$\frac{t}{kg}$	computed	2.92638 5.03132	t km	impact DB	0.842	$\frac{kg\ co2e}{t\ km}$	computed	2.464 4.236

$\{o1, f1, g1, w1\}$ are fully connected, as they pairwise share event $ev2$. Object graphs are used as a basis for multiple OCPM techniques intended to connect objects and/or events for specific considerations, such as identifying single process executions [1], [13]. A version of such a graph can be used to allocate impacts. For example, to allocate the impacts of the registration event *ev1* to one of the two bikes, we could determine the shortest path of the object involved in the event (*o1*) to each of the bikes and assign the impacts to the object with the shortest path. If the shortest path leads to multiple targets, we split the impact according to a specified weighting. Allocating all impacts according to this method and splitting shared impacts (e.g. the impact of the packing installation) equally leads to a total impact of $236.335\ kq\ co2e$ for bike1 and $195.913\ kq\ co2e$ for bike2.

Based on the demonstration above, we observe that the proposed method allows for flows tied to objects, event-object combinations and events can be combined without additional pre-processing. It is even possible to include relevant data that is tied to a subset of objects by adding a dummy event with the flow/impact data that only has E2O relations to the objects of the subset. However, flows and flow parameters associated with objects must either be time-independent (no aggregation over a period of time) or span the entire timeframe the OCEL represents (e.g., a machine's energy consumption).

An important realisation is that the consideration of individual events and objects does **not** imply that more detailed sustainability data is required (but it can be used if available). Instead, it merely enables more accurate allocation. Consider the example of the consumed energy (flow 2) from Section II-B: The monthly energy consumption tied to the activity is equally distributed among the frames processed within the considered timeframe. Leveraging the OCEL, we do not distribute the consumption among the products but the logged events of the activity. Now, if one of the products required rework, i.e., the activity is executed multiple times, it receives a higher impact[7]. Another interesting realisation is that the multi-perspective consideration of the process and the allocation based on E2O relations allows for a mixture of aggregation levels w.r.t. quantified flows and impacts.

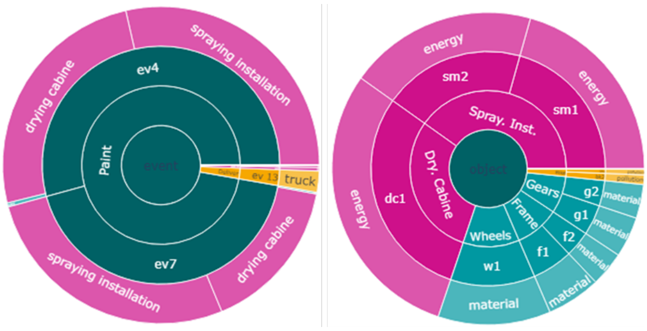
[7]Considering the products with a higher impact than others, an analyst can now identify the necessity of rework as an impact driver and apply PM techniques to determine the cause of the rework can now be identified using OCPM techniques and help avoid the additional impact in future.

Interpretation: Figure 6a shows the events and objects and the impacts they directly refer to, and the results of the allocation are depicted in Subfigure 6b. We see that analyses based on OCEL data allow for an interpretation of the results from an event and object perspective. On the one hand, the event perspective gives insights into which activities a company may need to change in order to improve its sustainability performance. On the other hand, the object perspective allows for identifying machinery, components, or products with high sustainability impacts. It is also possible to interpret the results on instance (e.g. individual components, bikes, ...) or type level (e.g. all bikes produced in the time frame represented in the event log). This allows for a joint consideration of LCA-concepts such as flows and impacts with business processes, enabling an alignment between sustainability and process management. This lays a technical foundation for the application of more advanced process analyses. For example, root-cause analysis [29] can be used to identify correlations between individual process behaviour, object properties, and sustainability impacts.

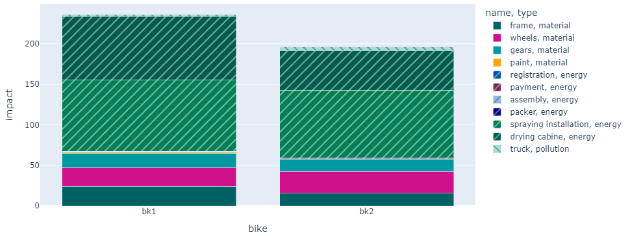
The comparison of the allocated impacts per bike shown in Figure 6b emphasises another interesting distinction to classical LCA: The allocation of impacts to the individual bikes differs according to the properties of the objects and events. In part, this is due to the more fine-grain data required for OCEL. However, the fact that we did not have to allocate individual flow parameters and work with averages also accounts for this phenomenom. This also offers the basis for a more accurate interpretation of results: The less allocation has to be done in the inventory analysis phase (e.g., by having to divide a monthly energy consumption by an average number of bikes produced), the less falsified and flattened the data we use for the detection of impact drivers.

V. Implementation: OCEAn

OCEAn (**O**bject-**C**entric **E**nvironmental **An**alysis) is intended as a proof of concept for researchers and practitioner's to explore OCELs and the application of OCPM for sustain-



(a) Impacts of flows directly connected to the different events and objects.



(b) All impacts allocated to the two different bikes based on their relationship with the origin of the impact in the object graph.

Figure 6: Overview of the impacts that can be derived based on the data in the event log.

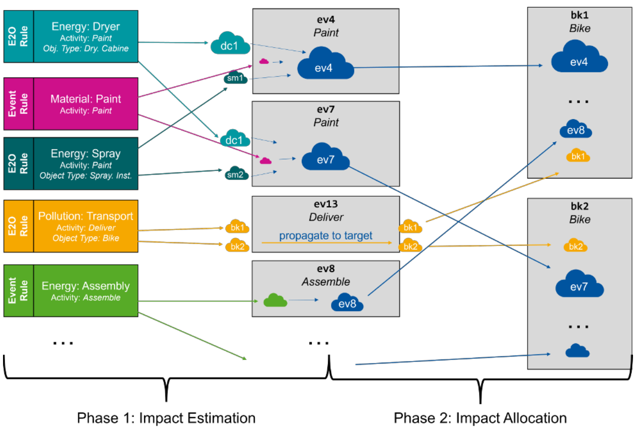


Figure 7: Overview of the functionality of OCEAn.

ability assessment and is publicly available on GitHub [8]. The tool currently only considers emissions, i.e., assessments for the impact category *climate change* and assumes all flows to be tied to either an event or an event-to-object relation. Object flows can easily be transformed to events by creating a dummy event for every object representing a flow which is related only to the object is was created for.

Figure 7 provides a high-level description of the functional-

[8]https://github.com/rwth-pads/ocel4lca, repository also includes the illustrative example and all calculations made with it.

Figure 8: Dashboard describing for the uploaded example OCEL. The emissions displayed in the activity overview at the bottom are based on the impacts determined using the impact rules.

ity of OCEAn. OCEAn extends an OCEL with impacts which are linked to either events or objects. It allows the user to define *impact rules* on activity and attribute level to quantify flows included in the event log. These event-related impacts are then either aggregated to a single impact value per event, or allocated to a selected object type. The resulting OCEL then contains an addition attribute stating either the event's or the object's impact.

A user can select an OCEL 2.0 as sqlite or browse through a selection of publicly available OCELs and load the event log. Figure 8 shows the user interface after loading an event log. The dashboard shows the user a sample of the uploaded event and object data and allows for the discovery of a process model (object-centric Petri net). The user gets an overview of the event and object attributes (different tabs), showing their names, the activities and object types they refer to, the number of unique values, and the range of numeric values, helping her specify each attribute's unit. Furthermore, she can select attributes that already quantify an impact and identify the individual object types as either a process resource or a handling unit; this distinction decreases the allocation's required computational power and memory. At the bottom of the page, the user can see the frequency of events for the different activities and objects per object type. Once impacts are quantified, she can visualise the average, sum, min, max and median impact per activity as well as the total number or share of events that have an impact assigned to them.

In the *Emissions*-Tab (see Figure 9), the user can define impact rules for either events or event-to-object combinations. She selects an activity (and object type), based on which she can choose the attribute's of which she would like to multiply the values with to quantify every event(-to-object)'s flow. If

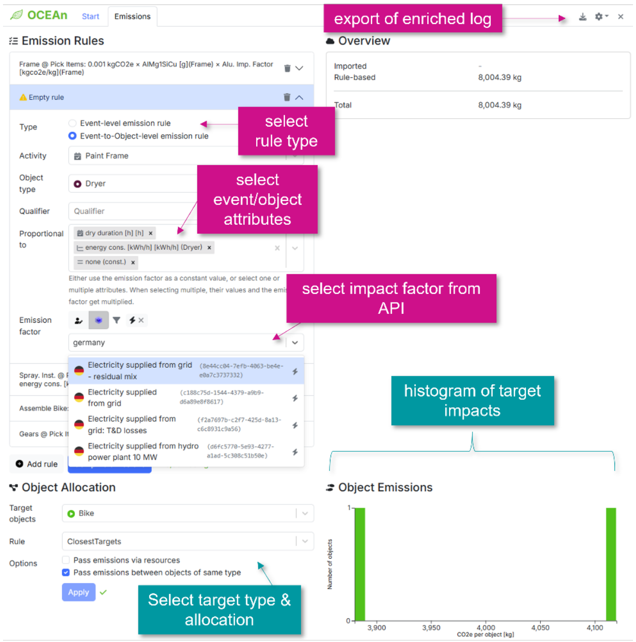


Figure 9: The emissions tab to define impact rules and allocate impacts to objects of a selected object type.

she adds a climatiq[9] API Key in the settings, she can search for a corresponding impact factor in the catalogue. OCEAn applies filters to the catalogue based on the combination of specified units to support the selection of a suitable impact factor. In the top-right, the user gets an overview of the total impacts determined, divided into the impacts provided directly in the event log and those computed by the rules.

Once OCEAn detects impacts for the OCEL, these impacts can be allocated to the objects of a selected object type. The rationale behind the allocation is illustrated by Figure 7. Every impact rule leads to the calculation of an impact for all events of the selected activity for the specified flow. The definition of multiple impact rules, therefore, may lead to several flow's impacts referring to the same event. The event-based allocation merely aggregates all impacts associated with an event and assigns this sum to the event. For the target-based allocation, OCEAn has three simple allocation rules the user can choose from. Each of the rules first assigns all E2O-impacts where the object is a target object to this target. The remaining impacts are aggregated to a single value per event. Each of the allocation rules assigns a set of target objects to every event and splits the remaining impact evenly among the event's set of objects. The allocation rules differ in their approach of specifying this set of targets. The ALLTARGETS allocation rule assigns all target objects to every event; all impacts are divided evenly among the target objects. In the PARTICIPATINGTARGETS allocation rule, we distinguish between events that have an event-to-object relation to at least one target and those who do not. For each event with such a relation, the assigned set of targets are all targets the event is related with. All other events map to all target objects; for these events we apply the ALLTARGETS-rule. Finally, for the CLOSESTTARGETS allocation rule, we determine every related object's shortest path to a target object in the object graph. Every event's remaining impacts are assigned to all targets that can be reached within a distance of the shortest path.

[9]Emission database we use: https://app.climatiq.io/. Every user gets keys with a limited number of calls for free.

VI. DISCUSSION

The work presented here is explorative; we demonstrate the suitability of OCEL for sustainability analyses using a realistic yet fictive example and show that corresponding software tools can be implemented. As seen in the previous sections, it is possible to generate detailed impact profiles of individual events, objects, activities, object types, process executions and individual processes based on an OCEL. To turn this demonstration into fruitful discourse, we consider the suitability, feasibility, challenges and potentials of using OCPM for sustainability analysis.

In its structure, an OCEL is comparable to data formats used in LCA, such as EcoSpold 2 [31] and ILCD [12]; sustainability-relevant data can be stored and processed from an OCEL. There are, however, notable conceptual differences between PM and LCA. An event log describes a discrete world-view, while in LCA, traditionally, continuous processes are modelled. OCEL-based LCA has the potential to align the sustainability perspective with the business perspective, as it builds upon traces of business processes and not on manually created models. It can be used to identify root causes for sustainability impacts and give insights into what needs to be changed in the business processes of an organisation. The explicit hierarchy inherently provided by the structure of the OCEL (events/objects/attribute values vs. activities/object types/attributes) allows for results on different levels of aggregation and a more differentiated analysis. As we have shown, it is possible to conduct flexible analyses from multiple perspectives with our approach, which addresses the need for more flexible analysis possibilities in LCA. However, sustainability analysis based on OCELs has one major limitation compared with classical LCAs: Everything that can be considered must be either an event or an object – both of which require unique identifiers. Yet, the assumption that in a production and logistics setting every element of interest and relevance in process and sustainability analysis (e.g., individual final products such as pens or cups) have an identifier is unrealistic. This being said, there is ongoing research in the domain of PM that looks into the consideration of non-identifiable products that play a role in business processes [17]. We have seen above that the OCEL can be used to both store and determine data relevant for sustainability analysis. However, the determination of relevant data using an OCEL strongly relies on domain knowledge and a more detailed understanding of the data, i.e. aitomation is challenging. On the other side, with a well-defined data standard for sustainability assessment it is conceivable that OCPM offers a foundation for continuous automated assessment—especially in combination with impact databases.

The creation of an OCEL and the collection of holistic sustainability data are challenging tasks which both facing similar challenges such as the combination of data from different systems (potentially crossing organisational boundaries). Not all data required for sustainability assessment is native to the data required for an OCEL and the share of data that has to be added specifically for sustainability assessment is large and cannot be neglected. However, we demonstrated that logs from operational systems (event data) and master data (e.g., weights, distances, (amount of) materials used) can be beneficial for sustainability analysis and even lead to more accurate results. The question of how to calculate and allocate impacts to different entities needs further elaboration in the future. We have addressed this in the OCEAn implementation by showcasing different variants of allocation rules. In the future, these can be elaborated e.g. with introducing allocation rules that consider weights or economic value of entities, as is common in LCA.

Finally, our approach demonstrates that the distinction of events and objects in OCPM is well suited to integrate sustainability data and inform sustainability analyses. However, OCPM is a relatively young research area, and challenges remain, e.g. in the generation of process models based on OCEL that support the visualisation and interpretation of sustainability impacts. This is an ongoing research area.

VII. Conclusion

In this paper, we explored the potential of object-centric process mining (OCPM) as a novel approach to sustainability analysis. Our work demonstrates the promising capabilities of OCPM in supporting a semi-automated multi-perspective sustainability analysis by combining domain knowledge with object-centric event logs. Our approach addresses several critical limitations of traditional sustainability assessment methods. By leveraging existing digital traces of business activities, the outlined approach could reduce the effort required for manual data collection. The framework provides flexibility in data integration, allowing iterative enrichment of event logs with sustainability-relevant information. The key contribution of this research is the ability to align sustainability perspectives directly with business processes. By building upon actual traces of organizational activities, our method offers unprecedented insights into the root causes of environmental impacts. Using OCELs enables organisations to not just measure their sustainability performance, but to understand and potentially modify the underlying processes that generate these impacts using object-centric process mining techniques. Based on the findings in this work, we consider sustainability analysis using OCPM a promising new direction worth pursuing. Hence, future work should focus on developing a more concrete conceptualisation of object-centric process mining for sustainability, and focus on:

- The consideration and integration of relevant process elements without unique identifiers
- Developing more sophisticated allocation rules for impact assessment
- Integrating domain-specific knowledge more comprehensively, e.g., with a standard for sustainbility-enhanced OCELs (sOCEL)
- Exploring methods to natively log sustainability-relevant data in information systems to leverage the more detailed perspective enabled with OCPM
- Advancing visualization techniques for interpreting sustainability impacts in object-centric process models

While not a complete solution, our framework represents a significant step towards more dynamic, data and process driven sustainability analysis. By bridging the gap between process mining and life cycle assessment, we offer a foundation for more agile and insightful impact evaluation. As organisations increasingly seek to understand and reduce their environmental footprint, techniques based on OCPM will become crucial. The potential to transform raw business process data into meaningful sustainability insights offers a powerful tool for driving organizational change towards more sustainable practices.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC-2023 Internet of Production - 390621612.

References

[1] Jan Niklas Adams, Daniel Schuster, Seth Schmitz, Gunther Schuh, and Wil M.P. van der Aalst. Defining Cases and Variants for Object-Centric Event Data. In *2022 4th International Conference on Process Mining (ICPM)*, pages 128–135, Bolzano, Italy, October 2022. IEEE.

[2] Martin Baitz, Stefan Albrecht, Eloise Brauner, Clare Broadbent, Guy Castellan, Pierre Conrath, James Fava, Matthias Finkbeiner, Matthias Fischer, Pere Fullana I Palmer, Stephan Krinke, Christian Leroy, Oliver Loebel, Phil McKeown, Ivo Mersiowsky, Bernhard Möginger, Marcus Pfaadt, Gerald Rebitzer, Elmar Rother, Klaus Ruhland, Aafko Schanssema, and Ladji Tikana. LCA's theory and practice: Like ebony and ivory living in perfect harmony? *The International Journal of Life Cycle Assessment*, 18(1):5–13, January 2013.

[3] Alessandro Berti, Istvan Koren, Jan Niklas Adams, Gyunam Park, Benedikt Knopp, Nina Graves, Majid Rafiei, Lukas Liß, Leah Tacke Genannt Unterberg, Yisong Zhang, Christopher Schwanen, Marco Pegoraro, and Wil M. P. van der Aalst. OCEL (Object-Centric Event Log) 2.0 Specification, 2024.

[4] Alessandro Berti, Sebastiaan van Zelst, and Daniel Schuster. PM4Py: A process mining library for Python. *Software Impacts*, 17:100556, September 2023.

[5] Anders Bjørn, Mikołaj Owsianiak, Christine Molin, and Michael Zwicky Hauschild. LCA History. In *Life Cycle Assessment*. Springer, 2018.

[6] Lars Brehm, Jessica Slamka, and Andreas Nickmann. Process Mining for Carbon Accounting: An Analysis of Requirements and Potentials. In Jens Dibbern, Jens Förderer, Thomas Kude, Franz Rothlauf, and Kai Spohrer, editors, *Digitalization Across Organizational Levels*, pages 209–244. Springer International Publishing, Cham, 2022.

[7] Felipe Cerdas, Sebastian Thiede, Max Juraschek, Artem Turetskyy, and Christoph Herrmann. Shop-floor Life Cycle Assessment. *Procedia CIRP*, 61:393–398, 2017.

[8] Dries Couckuyt and Amy Van Looy. Green BPM as a business-oriented discipline: A systematic mapping study and research agenda. *Sustainability*, 11(15), 2019.

[9] Dries Couckuyt and Amy Van Looy. A systematic review of green business process management. *Business Process Management Journal*, 26(2):421–446, 2019.

[10] John Elkington. Towards the sustainable corporation: Win-win-win business strategies for sustainable development. *California Management Review*, 36(2):90–100, 1994.

[11] John Elkington. 25 years ago I coined the phrase “Triple bottom line.” here’s why it’s time to rethink it. *Harvard Business Review*, 2018.

[12] European Commission, Joint Research Centre, and Institute for Environment and Sustainability. *International Reference Life Cycle Data System (ILCD) Handbook :General Guide for Life Cycle Assessment : Detailed Guidance.* Publications Office, LU, 2010.

[13] Dirk Fahland. Process Mining over Multiple Behavioral Dimensions with Event Knowledge Graphs. In Wil M. P. Van Der Aalst and Josep Carmona, editors, *Process Mining Handbook*, volume 448, pages 274–319. Springer International Publishing, Cham, 2022.

[14] Andreas Fritsch. Sustainability Analysis Patterns for Process Mining and Process Modelling Approaches. In *ICPM 2024 International Workshops*, 2024.

[15] Andreas Fritsch, Johanna Hammerstein, Clemens Schreiber, Stefanie Betz, and Andreas Oberweis. Pathways to Greener Pastures: Research Opportunities to Integrate Life Cycle Assessment and Sustainable Business Process Management Based on a Systematic Tertiary Literature Review. *Sustainability*, 14:11164, 2022.

[16] Deloitte Global. *Global Process Mining Survey 2025 - Delivering Value with Process Mining – Adoption and Success Factors of Process Mining*. April 2025.

[17] Nina Graves, Tobias Brockhoff, István Koren, and Wil M. P. van der Aalst. Extending process intelligence with quantity-related process mining. In Andrea Delgado and Tijs Slaats, editors, *Process mining workshops*, page 337–349, Cham, 2025. Springer Nature Switzerland.

[18] Nina Graves, István Koren, and Wil M. P. van der Aalst. ReThink Your Processes! A Review of Process Mining for Sustainability. In *The 9th International Conference on ICT for Sustainability (ICT4S 2023)*. IEEE, 2023.

[19] Green Software Foundation. Introduction to the impact framework, 2025. Accessed: 2025-05-01.

[20] Jeroen B. Guinée, Reinout Heijungs, Gjalt Huppes, Alessandra Zamagni, Paolo Masoni, Roberto Buonamici, Tomas Ekvall, and Tomas Rydberg. Life Cycle Assessment: Past, Present, and Future. *Environmental Science & Technology*, 45(1):90–96, 2011.

[21] Paul Harmon. The scope and evolution of business process management. In *Handbook on Business Process Management 1*. Springer, 2015.

[22] Patrick Haun, Philipp Müller, and Marzia Traverso. Improving automated life cycle assessment with life cycle inventory model constructs. *Journal of Cleaner Production*, 370:133452, 2022.

[23] Eva Heiskanen. Managers’ interpretations of LCA: Enlightenment and responsibility or confusion and denial? *Business Strategy and the Environment*, 9(4):239–254, 2000.

[24] Eva Heiskanen. The institutional logic of life cycle thinking. *Journal of Cleaner Production*, 10(5):427–437, 2002.

[25] Ourania Areta Hiziroglu and Onur Dogan. Sustainability Performance Through a Business Process Mining Based Conceptual Framework Integrating GRI Metrics. In *Artificial Intelligence: Towards Sustainable Intelligence*. 2025.

[26] ISO. ISO 14040:2006 Environmental management - Life cycle assessment - Principles and framework, 2006.

[27] ISO. ISO/TS 14072:2014 Environmental management - Life cycle assessment - requirements and guidelines for organizational life cycle assessment, 2014.

[28] Finn Klessascheck, Ingo Weber, and Luise Pufahl. SOPA: A framework for sustainability-oriented process analysis and re-design in business process management. *Information Systems and e-Business Management*, January 2025.

[29] Benedikt Knopp, Mahsa Pourbafrani, and Wil M.P. van der Aalst. Root cause analysis using rule mining on object-centric event logs. In *Process Mining Workshops, ICPM 2024 International Workshops*, Copenhagen, Denmark, 2025. Springer.

[30] Istvan Koren, Niklas Adams, and Alessandro Berti. Ocel 2.0 resources – www.ocel-standard.org. 2024.

[31] Ingo Meinshausen, Peter Müller-Beilschmidt, and Tobias Viere. The EcoSpold 2 format—why a new format? *The International Journal of Life Cycle Assessment*, 21(9):1231–1235, 2016.

[32] Nigel P. Melville. Information Systems Innovation for Environmental Sustainability. *MIS Quarterly*, 34(1):1–21, 2010.

[33] Christian Ortmeier, Nadja Henningsen, Adrian Langer, Alexander Reiswich, Alexander Karl, and Christoph Herrmann. Framework for the integration of Process Mining into Life Cycle Assessment. *Procedia CIRP*, 98:163–168, 2021.

[34] Lars Reinkemeyer, editor. *Process Mining in Action: Principles, Use Cases and Outlook*. Springer International Publishing, Cham, 2020.

[35] Daniel Schneider, Patrick Jordan, Jonathan Dietz, Michael F. Zaeh, and Gunther Reinhart. Concept for Automated LCA of Manufacturing Processes. *Procedia CIRP*, 116:59–64, 2023.

[36] Wil M.P. van der Aalst. *Process Mining*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2016.

[37] Wil M.P. van der Aalst. Object-Centric Process Mining: Dealing with Divergence and Convergence in Event Data. In Peter Csaba Ölveczky and Gwen Salaün, editors, *Software Engineering and Formal Methods*, volume 11724 of *Lecture Notes in Computer Science*, pages 3–25. Springer, 2019.

[38] Wil M.P. van der Aalst and Ernesto Damiani. Processes Meet Big Data: Connecting Data Science with Process Science. *IEEE Transactions on Services Computing*, 8(6):810–819, 2015.

[39] WBCSD and WRI. The Greenhouse Gas Protocol, 2004.