Twin Transitions Powered By Event Data

Using Object-Centric Process Mining To Make Processes Digital and Sustainable

Wil M.P. van der Aalst[1,2,3]

[1]Process and Data Science, RWTH Aachen University, D-52074 Aachen, Germany
[2]Celonis, Theresienstrasse 6, D-80333 München, Germany
[3]Fraunhofer-Institut für Angewandte Informationstechnik, Schloss Birlinghoven, D-53757, Sankt Augustin,Germany

Abstract
Digital transformation plays a major role in realizing sustainability goals. Inefficiencies and a lack of transparency lead to excessive resource usage and carbon emissions. For example, containers are idle over half their lifetime and often shipped empty. Process mining helps to uncover inefficiencies and create transparency. Process discovery can be used to show the actual processes and conformance checking can be used to diagnose deviations. However, traditional process mining focuses on individual cases. To improve sustainability, one needs a more holistic view considering all objects involved. Optimizing the process by focusing on a single object type may lead to unnecessary waste and emissions. Therefore, this keynote papers advocates using *Object-Centric Process Mining* (OCPM). OCPM considers all objects, events, and their interrelations. OCPM views activities from any perspective using a *single source of truth*. Data are extracted only once and better reflect the true operations, thus supporting digital transformation. Having a reliable view of the actual operations and their environmental impact is the first step toward avoiding the long-term depletion of natural resources while addressing current needs. Moreover, opportunities to improve sustainability live at the intersections of events and objects. OCPM helps to uncover these opportunities. Therefore, this new form of process mining enables twin transitions that make processes both green and digital.

Keywords
process mining, digital transformation, sustainability, Petri nets, object-centric process mining

1. Introduction

The term *twin transition* emerged in the context of major European policy initiatives such as the *European Green Deal* (presented in 2019 and approved in 2020 with the goal of making the EU environmentally sustainable) and the *European Digital Strategy* (consisting of various programs, including the Digital Decade aimed and realizing digital transformations in four areas: skills, infrastructure, business, and government). Combining digital transformations and the realization of sustainability goals makes perfect sense. In this keynote paper, we focus on the role of *Object-Centric Process Mining (OCPM) in realizing twin transitions in operational processes.* Figure 1 shows how these topics are related.

Algorithms & Theories for the Analysis of Event Data and Petri Nets for Twin Transition 2023, 26-27 June 2023, Caparica, Lisbon, Portugal
wvdaalst@pads.rwth-aachen.de (W. M.P. van der Aalst)
https://vdaalst.com/ (W. M.P. van der Aalst)
0000-0002-0955-6940 (W. M.P. van der Aalst)

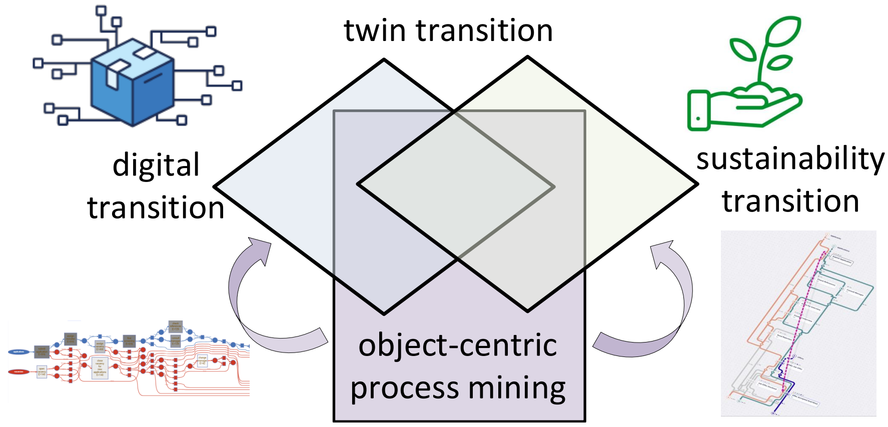

Figure 1: Twin transition(s) combine digital transition(s) and sustainability transition(s). Object-Centric Process Mining (OCPM) plays an enabling role in this.

The traditional focus of process mining has been on *performance* (e.g., minimizing costs and time, maximizing profit and productivity) and *compliance* (e.g., detecting fraud and risk) to realize top-line growth and bottom-line growth. The top line refers to a company's revenues or gross sales. The bottom line refers to a company's net income. Now the so-called "green-line" has been added by making sustainability a new priority. Obviously, improving the green-line does not imply lowering the top-line and bottom-line growth. Given the right incentives, minimizing the costs may imply minimizing carbon emissions. Consider, for example, the shipping of containers. It is estimated that 80-90% of the world's goods are transported in ocean freight shipping containers. However, many containers are shipped empty (estimates range from 30%-60%). Reducing the number of containers shipped empty will lower costs and emissions. Clearly, information systems play a key role in this [1].

The synergistic effects between reducing costs and carbon emissions are evident. Therefore, it is no surprise that process mining vendors like *Celonis* (www.celonis.com) can play a major role in realizing twin transitions. Celonis developed apps running on top of the process-mining platform to reduce carbon emissions, save time, cut costs and improve sustainability across supply chains. These applications were developed in collaboration with organizations like *Climatiq* (www.climatiq.io), *EcoVadis* (www.ecovadis.com), and *IntegrityNext* (www.integritynext.com) that specialize in quantifying sustainability factors like the emission of greenhouse gases. The Climatiq Carbon Calculation Engine can, for example, be used to estimate emissions related to shipments and other logistics activities.

However, such improvements are only possible when events are recorded properly. Many events relevant to sustainability are not recorded at all or are extremely difficult to extract from existing systems. CRM (Customer Relationship Management) and ERP (Enterprise Resource Planning) systems from vendors like SAP, Oracle, and Salesforce have thousands of tables with low-level events stored in an incomprehensible manner. Also, systems of different vendors use different names and table structures. This makes things incomparable. It is not realistic to assume that these systems agree on common formats and naming conventions. However, high-end process-mining tools like Celonis can be seen as a layer on top of these systems providing a unified view at the level of true business operations. This is where *Object-Centric Process Mining* (OCPM) [2] comes into play. OCPM facilitates digital transformations and the data collected can be used to measure and realize sustainability goals [3].

2. Object-Centric Process Mining (OCPM)

Process mining techniques use event data and process models (e.g., expressed as a Petri net, a DFG, or BPMN model) to improve processes. It is possible to *discover* process models based on event data. These discovery approaches can be grouped into *bottom-up process discovery approaches* like the Alpha algorithm and region-based techniques [4, 5, 6, 7, 8, 9, 10] and *top-down process discovery approaches* like the inductive mining approaches [11, 12, 13]. See [14] for a recent survey of process discovery techniques. It is also possible to *check* the *conformance* of processes by comparing a process model with event data. The two most frequently used conformance-checking approaches are *token-based replay* [15] and *alignments* [16, 17]. Since each event has a timestamp, it is easy to analyze performance by replaying event data on models. There are also process mining approaches to predict the evolution of the whole process or individual process instances. By connecting process mining to low-code automation platforms, it is also possible to automatically perform actions addressing performance and compliance problems. Many of the more advanced process mining techniques have their roots in Petri nets, because real-life processes are inherently concurrent. For some time, process mining was seen as just another"academic exercise" having to practical impact. However, now it has become one of the leading approaches in *Business Process Management* (BPM) and is widely applied. According to Gartner, there are around 40 process mining vendors (also see www.processmining.org), and process mining now has its own Gartner "Magic Quadrant", showing that it has become a standard product category [18].

Traditional approaches for process modeling and process analysis tend to focus on one type of objects (also referred to as cases or instances) and each event refers to precisely one such object. However, real-life activities may involve *multiple objects* of *different types*. These objects may be related. However, in traditional process mining, once needs to pick a case notion and it is assumed that cases do not interact. Moreover, data extraction is time-consuming and needs to be repeated when new questions emerge. *Object-Centric Process Mining* (OCPM) [19, 2] techniques do not enforce a case notion. Each event may refer to any number of objects, possibly of different types, instead of a single case identifier. As a result, it is possible to view events and objects from any perspective using a *single source of truth*. There have been several other proposals using artifact-centric models [20, 21], object-centric behavioral constraint models [22, 23], multi-perspective models [24], multi-event logs (Celonis), eXtensible Object-Centric (XOC) event logs [25], graph databases tailored towards event data [26], catalog and object-aware nets [27], etc. However, these approaches turned out to be too complex or not scalable. This is why we resorted to the simpler *Object-Centric Process Mining* (OCPM) a few years ago and others are following now. For OCPM, we make the following assumptions:

- Each *object* has precisely one *object type*, but many objects may have the same type.

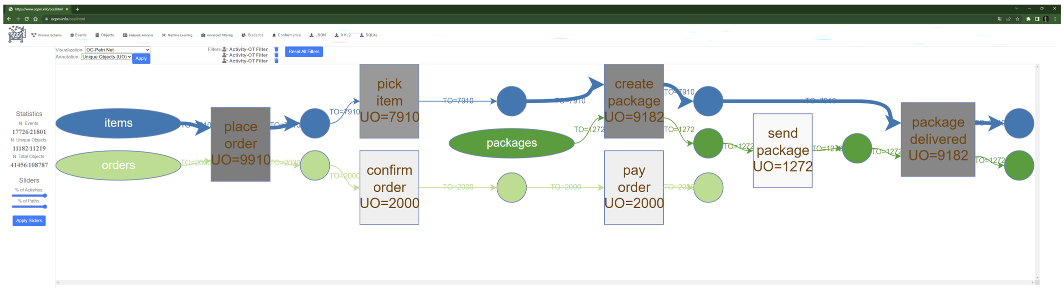

Figure 2: Object-Centric Petri Net (OCPN) discovered using the OC-PM tool (ocpm.info).

- Each *event* has an *event type*, also called *activity*. Many events can have the same type, but each event has precisely one type.
- An event refers to *any number of objects* using so-called *Event-to-Object* (E2O) relations.
- Objects can be related to other objects using *Object-to-Object* (O2O) relationships. These relations are static.
- Both objects and events can have any number of *attributes* with corresponding *values*. Object attribute values have a timestamp and event attribute values do *not* have a timestamp.

In the PADS group, we developed several non-commercial open-source tools supporting object-centricity, e.g., the "OCELStandard" package in ProM (promtools.org), the OC-PM tool (ocpm.info), and Object-Centric Process Insights (ocpi.ai). These ideas were also implemented in the Celonis, the leading commercial process mining tool. The Celonis Process Sphere implementation was presented during Celosphere 2022 [28].

Figure 2 shows a very simple Object-Centric Petri Net (OCPN) discovered from Object-Centric Event Data (OCED) using the OC-PM tool (implementing the approach described in [2]). One can think of OCPN as an underspecified Colored Petri Net (CPN) [29]. The different colors refer to the three object types: *orders* (light green), *items* (blue), and *packages* (dark green). Note that one order may refer to a variable number of items. A package may contain a variable number of items originating from different orders. Items from one order can be distributed over multiple packages. This cannot be captured in a workflow net (Petri net with a start and end place), a Directly-Follows Graph (DFG), or a BPMN model, because these describe one case in isolation. Note that an OCPN is deliberately underspecified because the relations between orders, items, and packages are in the data and cannot be computed using arc expressions like in CPN Tools. See [19, 28, 2] for more information about Object-Centric Process Mining (OCPM).

How does this relate to sustainability and twin transitions? First of all, OCPM makes it easier to store data in a system-independent manner. Operational processes tend to leave data in different systems. OCPM provides a layer on top of these systems unifying the data in a format much closer to the actual business operations. Instead of storing the data in a format close to the source system (e.g., SAP with its 800.000 tables), data should be stored in a well-defined unified format using names for object types and event types that end-users understand. This is vital for digital transformation. Second, the resulting *single source of truth* can be used to track and improve sustainability goals. Objects may refer to products, resources, emission rights,

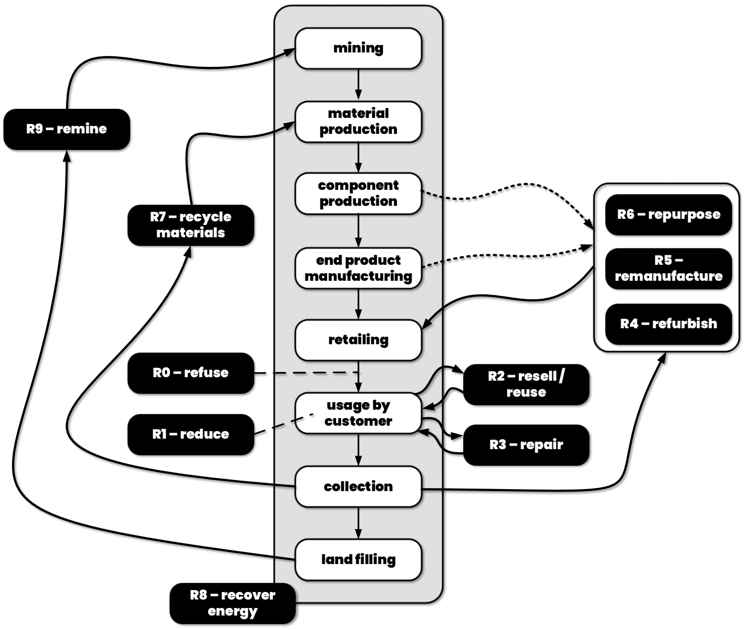

Figure 3: The ten resource value retention options, also known as the ten R-imperatives [30], showing different ways the get closer to a Circular Economy (CE) [1].

waste, etc. Events describe all relevant business operations and it is possible to enrich these with information about emissions. This allows quantifying and monitoring sustainability aspects of operational processes, thus providing end-to-end transparency. This increased transparency also helps to make better decisions, e.g., select suppliers and means of transport.

One can view sustainability as just another Key Performance Indicator (KPI) next to profit margin, average order fulfillment time, customer churn rate, conversion rate, etc. However, it is far from trivial to define sustainability KPIs. Organizations like Climatiq, EcoVadis, and IntegrityNext specialize in quantifying sustainability factors. Using OCPM, it is possible to address this in a holistic manner, e.g., computing the carbon footprint of a product or service taking into account all process steps.

3. 10R: Refuse, Reduce, Resell/Reuse, Repair, Refurbish, Remanufacture, Repurpose, Recycle, Recover, or Remine?

Sustainability considerations will lead to new trade-offs requiring decision support. Twin transitions require fundamentally rethinking how products and services are created and delivered. To illustrate this, we focus on the well-known *Resource Value Retention Options*, better known as the *R-imperatives* [30]. Figure 3 shows ten R-imperatives [1].

R-imperative R0 (Refuse) refers to not having the product or using the service in the first place. R1 (Reduce) refers to using products less frequently or using them with more care, such that they can be used longer. R2 (Resell/Reuse) refers to the usage by a second, third, fourth, etc. customer using the product without any modifications. R-imperative R3 (Repair) refers to extending the lifetime of the product without making changes to its function. R4 (Refurbish) refers to replacing components of a product without completely disassembling it. R5 (Remanufacture) refers to completely disassembling the product and check, clean, repair, or replace components. R6 (Repurpose) refers to the usage of a product for a different purpose than initially intended. R7 (Recycle) refers to the reuse of the material without retaining the original structure or function. R8 (Recover) refers to capturing energy from waste or using by-products from production processes. R9 (Remine) refers to scrapping valuable materials and items from landfills. See [30] for a detailed description of these R-imperatives.

Whereas the R-imperatives refer to rethinking the whole product life-cycle, there are many less impactful decisions that can be made to improve sustainability, e.g., selecting a supplier or determining the mode of transport. Consider, for example, a materials-handling system in a car factory in Stuttgart that no longer functions due to a broken component (e.g., the engine to move the belt). It is possible to get a new component from the factory in Hamburg by rail or truck, the component can be refurnished in Stuttgart but this will take an additional week, and it is possible to reuse a component located in Madrid. The different possibilities have different costs and different ecological footprints. Note that the downtime of the materials-handling system itself also has sustainability effects. The example shows that such trade-offs require a wide variety of inputs and cannot be solved using a single viewpoint. Life-cycle sustainability assessment measures the impacts on the environment associated with the life cycle of a product or service. This is very difficult. For example, when producing a product, it is not clear how it will be used in the future (e.g., it is repurposed in a way not imagined during production). To make such life-cycle assessments less ad-hoc, we advocate the use of OCPM.

4. Outlook

In this keynote paper, we connected twin transitions with Object-Centric Process Mining (OCPM). Organizations struggle to realize digital transformations. Processes are still paper-based or data are stored in a heterogeneous landscape of information systems disconnected from the actual business operations. OCPM starts from system-independent data describing the most relevant events and objects. Such data can be extended with sustainability aspects allowing OCPM techniques to make trade-offs balancing costs, time, and environment.

OCPM provides a semantically enriched layer on top of existing systems to observe and improve operational processes. Because OCPM does not aim to replace existing information systems, it can be introduced gradually, focusing on the key pain points of an organization first. Systematically collecting information on events and objects allows for more transparency of real-life intra- and inter-organizational processes. This makes it easier to reliably quantify and reach sustainability goals. Combining OCPM with sustainability measures can uncover many execution gaps that also negatively impact the environment.

Acknowledgments

The author thanks the Alexander von Humboldt (AvH) Stiftung for supporting his research. The research is also funded by the Deutsche Forschungsgemeinschaft (DFG) under Germany's Excellence Strategy, Internet of Production (390621612).

References

[1] W. van der Aalst, O. Hinz, C. Weinhardt, Sustainable Systems Engineering, Business and Information Systems Engineering 65 (2023) 1–6.

[2] W. van der Aalst, A. Berti, Discovering Object-Centric Petri Nets, Fundamenta Informaticae 175 (2020) 1–40.

[3] N. Graves, I. Koren, W. van der Aalst, ReThink Your Processes! A Review of Process Mining for Sustainability, 2023.

[4] W. van der Aalst, Process Mining: Data Science in Action, Springer-Verlag, Berlin, 2016.

[5] W. van der Aalst, A. Weijters, L. Maruster, Workflow Mining: Discovering Process Models from Event Logs, IEEE Transactions on Knowledge and Data Engineering 16 (2004) 1128–1142.

[6] A. Augusto, R. Conforti, M. Marlon, M. La Rosa, A. Polyvyanyy, Split Miner: Automated Discovery of Accurate and Simple Business Process Models from Event Logs, Knowledge Information Systems 59 (2019) 251–284.

[7] R. Bergenthum, J. Desel, R. Lorenz, S. Mauser, Process Mining Based on Regions of Languages, in: G. Alonso, P. Dadam, M. Rosemann (Eds.), International Conference on Business Process Management (BPM 2007), volume 4714 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, 2007, pp. 375–383.

[8] J. Carmona, J. Cortadella, M. Kishinevsky, A Region-Based Algorithm for Discovering Petri Nets from Event Logs, in: Business Process Management (BPM 2008), 2008, pp. 358–373.

[9] M. Solé, J. Carmona, Process Mining from a Basis of State Regions, in: J. Lilius, W. Penczek (Eds.), Applications and Theory of Petri Nets 2010, volume 6128 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, 2010, pp. 226–245.

[10] J. Werf, B. van Dongen, C. Hurkens, A. Serebrenik, Process Discovery using Integer Linear Programming, Fundamenta Informaticae 94 (2010) 387–412.

[11] S. Leemans, D. Fahland, W. van der Aalst, Discovering Block-structured Process Models from Event Logs: A Constructive Approach, in: J. Colom, J. Desel (Eds.), Applications and Theory of Petri Nets 2013, volume 7927 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, 2013, pp. 311–329.

[12] S. Leemans, D. Fahland, W. van der Aalst, Discovering Block-Structured Process Models from Event Logs Containing Infrequent Behaviour, in: N. Lohmann, M. Song, P. Wohed (Eds.), Business Process Management Workshops, International Workshop on Business Process Intelligence (BPI 2013), volume 171 of *Lecture Notes in Business Information Processing*, Springer-Verlag, Berlin, 2014, pp. 66–78.

[13] S. Leemans, D. Fahland, W. van der Aalst, Scalable Process Discovery and Confor-

mance Checking, Software and Systems Modeling 17 (2018) 599–631. doi:10.1007/s10270-016-0545-x.
[14] A. Augusto, R. Conforti, M. Dumas, M. Rosa, F. Maggi, A. Marrella, M. Mecella, A. Soo, Automated Discovery of Process Models from Event Logs: Review and Benchmark, IEEE Transactions on Knowledge and Data Engineering 31 (2019) 686–705.
[15] A. Rozinat, W. van der Aalst, Conformance Checking of Processes Based on Monitoring Real Behavior, Information Systems 33 (2008) 64–95.
[16] W. van der Aalst, A. Adriansyah, B. van Dongen, Replaying History on Process Models for Conformance Checking and Performance Analysis, WIREs Data Mining and Knowledge Discovery 2 (2012) 182–192.
[17] J. Carmona, B. van Dongen, A. Solti, M. Weidlich, Conformance Checking: Relating Processes and Models, Springer-Verlag, Berlin, 2018.
[18] M. Kerremans, K. Iijima, A. Sachelarescu, N. Duffy, D. Sugden, Magic Quadrant for Process Mining Tools, Gartner Research Note GG00774746, 2013. www.gartner.com.
[19] W. van der Aalst, Object-Centric Process Mining: Dealing With Divergence and Convergence in Event Data, in: P. Ölveczky, G. Salaün (Eds.), Software Engineering and Formal Methods (SEFM 2019), volume 11724 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, 2019, pp. 3–25.
[20] D. Fahland, Describing Behavior of Processes with Many-to-Many Interactions, in: S. Donatelli, S. Haar (Eds.), Applications and Theory of Petri Nets 2019, volume 11522 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, 2019, pp. 3–24.
[21] X. Lu, M. Nagelkerke, D. Wiel, D. Fahland, Discovering Interacting Artifacts from ERP Systems, IEEE Transactions on Services Computing 8 (2015) 861–873.
[22] G. Li, R. Medeiros de Carvalho, W. van der Aalst, Automatic Discovery of Object-Centric Behavioral Constraint Models, in: W. Abramowicz (Ed.), Business Information Systems (BIS 2017), volume 288 of *Lecture Notes in Business Information Processing*, Springer-Verlag, Berlin, 2017, pp. 43–58.
[23] A. Artale, A. Kovtunova, M. Montali, W. van der Aalst, Modeling and Reasoning over Declarative Data-Aware Processes with Object-Centric Behavioral Constraints, in: T. Hildebrandt, B. van Dongen, M. Röglinger, J. Mendling (Eds.), International Conference on Business Process Management (BPM 2019), volume 11675 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, 2019, pp. 139–156.
[24] M. Eck, N. Sidorova, W. van der Aalst, Discovering and Exploring State-Based Models for Multi-perspective Processes, in: M. Rosa, P. Loos, O. Pastor (Eds.), International Conference on Business Process Management (BPM 2016), volume 9850 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, 2016, pp. 142–157.
[25] G. Li, E. G. L. de Murillas, R. M. de Carvalho, W. van der Aalst, Extracting Object-Centric Event Logs to Support Process Mining on Databases, in: J. Mendling, H. Mouratidis (Eds.), Information Systems in the Big Data Era, CAiSE Forum 2018, volume 317 of *Lecture Notes in Business Information Processing*, Springer-Verlag, Berlin, 2018, pp. 182–199.
[26] S. Esser, D. Fahland, Multi-Dimensional Event Data in Graph Databases, Journal on Data Semantics 10 (2021) 109–141.
[27] S. Ghilardi, A. Gianola, M. Montali, A. Rivkin, Petri Net-Based Object-Centric Processes With Read-Only Data, Information Systems 107 (2022) 102011.

[28] W. van der Aalst, Object-Centric Process Mining: The Next Frontier in Business Performance, 2023. celon.is/OCPM-Whitepaper.
[29] K. Jensen, Coloured Petri Nets. Basic Concepts, Analysis Methods and Practical Use, EATCS monographs on Theoretical Computer Science, Springer-Verlag, Berlin, 1992.
[30] D. Reike, W. Vermeulen, S. Witjes, The Circular Economy: New or refurbished as CE 3.0? — exploring controversies in the conceptualization of the circular economy through a focus on history and resource value retention options, Resources, Conservation and Recycling 135 (2018) 246–264.