How Object-Centric Process Mining Helps to Unleash Predictive and Generative AI

Wil van der Aalst

professor at RWTH Aachen University and chief scientist at Celonis

Abstract:

Process mining has emerged as a pivotal discipline that bridges the gap between process science and data science, evolving significantly since its inception in the late 1990s. The discipline of process mining has been instrumental in addressing fundamental questions about actual vs. assumed processes, identifying bottlenecks and deviations, and predicting performance and conformance problems. Despite advancements in process discovery, conformance checking, and data-driven simulation, (1) data extraction remains challenging, (2) traditional case-driven approaches fail to identify problems involving multiple organizational units and processes, and (3) organizations fail to reap the benefits of the rapid developments in Artificial Intelligence (AI). The introduction of Object-Centric Process Mining (OCPM) and the integration with predictive and generative AI represent a revolutionary shift in process management. OCPM allows for a more nuanced analysis of processes without the constraints of a single-case notion, enabling a deeper understanding of the interactions between different object types within processes. This evolution towards a more faithful view of operational processes is further enhanced by the capabilities of predictive and generative AI, offering new opportunities for diagnosing and addressing operational problems. Next to an integration of OCPM and Predictive and Generative AI, we advocate a domain-specific approach to process mining. Leveraging standardized reference models powered by OCPM helps to accelerate the adoption of process mining.

(R)evolution of the Process Mining Discipline

Process mining, as we know it today, emerged as a novel discipline to bridge the gap between *process science* and *data science* [1]. Before the turn of the century, there were disjoint groups of scientists and practitioners. The first group, let's call them the "process scientists", focused on process management and process automation. The second group, the "data scientists", focused on transforming and analyzing data. The process scientists did not care much about data, and the primary focus was on modeling processes by hand and automating these processes using, for example, workflow management systems. The data scientists were not interested in processes and instead focused on supervised and unsupervised learning using tabular data, text, and images. At the time, this was mostly referred to as *Data Mining* (DM), but over time, *Machine Learning* (ML) and *Artificial Intelligence* (AI) became the more common terms to refer to this. Data-driven approaches (DM, ML, and AI) have dramatically improved over the last 25 years. This is visible to all, and today's process scientists are well aware of the fact that they need to use data. Moreover, data scientists increasingly realize that processes are key to the success of any organization. In the last two years, *Large Language Model* (LLM) chatbots such as ChatGPT, MS Copilot, and Bard, and text-to-image systems such as Stable Diffusion, Midjourney, and DALL-E served as a wake-up call for organizations. Also, higher management realizes that processes and the management of these processes will dramatically change. *Hybrid*

Intelligence (HI) will play a key role in the gradual redistribution of work between people and software/hardware [2]. However, organizations struggle to apply these new technologies to their operational processes, and process mining is still the only discipline that gives *equal attention* to both processes and data.

Figure 1 shows a timeline of process mining running from its inception in the late 1990-ties until now. What is interesting to note is that many of the original questions posed two decades ago are still valid [1,4]:

- What is the actual process, and how does it differ from the assumed or desired process?
- What are the main bottlenecks, and why are they there?
- What are the main compliance problems, and what do they have in common?
- Can we predict performance and conformance problems?
- What happens if we make this intervention?

Fifteen years ago (i.e., 2009), we already had a range of process discovery techniques, supported conformance checking and time prediction, and could generate full-fledged simulation models from event data [1]. Although process-mining techniques answering all of the above questions have been around for quite some time, the underlying problems are notoriously hard and still not fully solved. Seen from this perspective, we could argue that the process mining discipline is evolving steadily. However, as will be explained later, *Object-Centric Process Mining* (OCPM) and the amalgamation of process mining and *predictive and generative AI* can be viewed as a revolution rather than a steady evolution [3,4,5,6]. OCPM takes process mining to a new level where it is no longer needed to “straitjacket processes” using a single-case notion. Predictive and generative AI provides amazing possibilities, but in an enterprise setting, we need processes as the “lens” to look at the data. Together, they lead to a revolution in process management.

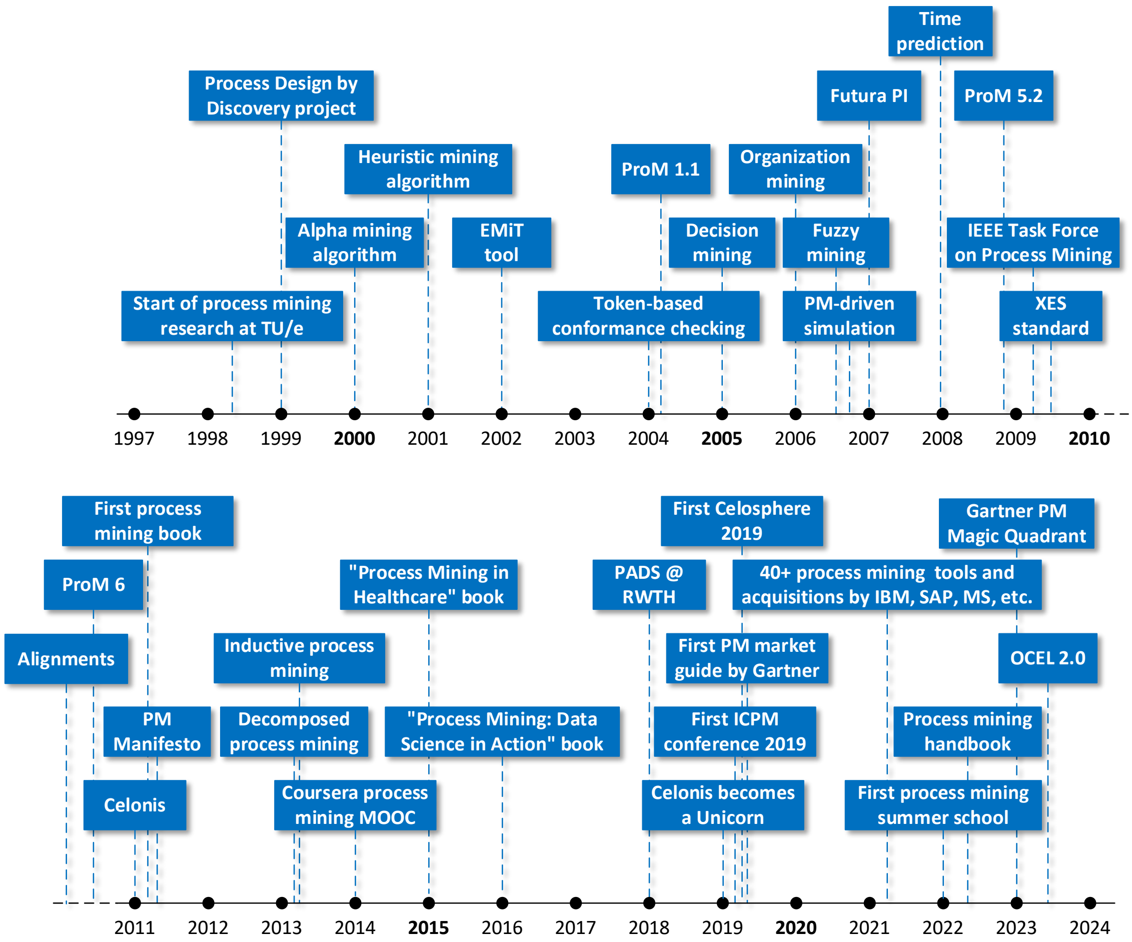

Figure 1: Timeline showing some of the main developments in process mining.

Figure 1 illustrates the increasing maturity of the process-mining discipline. On the academic side, process mining developed and is now a mature discipline. Half of the papers presented at leading BPM conferences, like the *International Conference on Business Process Management* (www.bpm-conference.org), are on process mining. The *International Conference on Process Mining* (www.icpmconference.org), established in 2019, attracts hundreds of process mining researchers each year. Based on the *Summer School on Process Mining* that took place in 2022, we created the Process Mining Handbook [4]. On the non-academic side, one can witness the availability of dozens of commercial process-mining offerings and an increasing adoption in organizations all over the globe [11]. This is reflected by the scale of the practitioner-oriented *Celosphere* conference (organized by Celonis) attracting thousands of participants. In 2023, Gartner also published the first *Magic Quadrant for Process Mining Tools* [9]. This proves that process mining can be viewed as a separate product category. In this book, you will find a wide range of successful applications of process mining. However, we are just at the beginning, and there are still challenges, as is discussed next.

Main Challenges and Opportunities

When applying process mining in the real world, one can witness the following challenges:

1. *Data extraction is involved* because information about objects (e.g., suppliers, orders, products, invoices, etc.) and events (i.e., activity executions involving these objects) is often scattered over multiple tables in different systems. One needs to locate these data and transform them into event data. This requires domain and technical expertise.

2. Traditional process model notations ranging from Business Process Modeling Notation (BPMN) to Directly-Follows Graphs (DFGs) make the implicit assumption that there is a *case notion*, i.e., the process model describes the life cycle of individual cases. This implies that *each process model represents a very specific view*. Changing the view requires adapting the case notion and going back to the source systems to extract new event data.
3. Although organizations tend to implement a range of highly similar processes (e.g., Accounts Payable, Accounts Receivable, Purchase to Pay, and Order to Cash), they use different IT systems that store data in different ways. Think of the table and column names in SAP (EKKO, EKPO, VBAK, VBAP, etc.) that are *system-specific*. Ideally, one would like to have a *system-agnostic single source of truth*. This enables comparative process mining focusing on the relevant differences over time and between organizations. Unifying data across organizations and systems is the only way to enable collaboration and share knowledge.
4. Most events involve *multiple objects* (e.g., a machine, an order, multiple components, and an operator). Yet, traditional case-centric process mining assumes one case per event. Therefore, *distortions* are introduced when enforcing a single-case notion. An event may be replicated for different objects, or causalities may get lost.
5. Operational problems often live at the intersections of processes and organizational entities. For example, a customer order may be delayed because of problems in procurement or production. Just looking at all events involving a delayed order will not reveal such problems. Therefore, one needs to understand the *interactions between the different object types*.
6. Users need to *understand* process mining results. It is easy to create ad-hoc dashboards showing Key Performance Indicators (KPIs). These may be easy to digest, but do *not* show the actual underlying processes. Using process discovery, one can reveal reality in an unbiased manner. However, showing a DFG may be overwhelming because of the *spaghetti-like* structure. Processes often have structures that are hidden using a simple DFG representation (especially when multiple objects are involved). Spaghetti-like DFGs are often caused by *concurrency* and *multiple objects*. If activities do not happen in a fixed order or involve multiple objects, DFGs will create lots of loops instead of showing the actual process structure.
7. Next to being understandable and showing the true fabric of processes, process mining results should also be *actionable*. This requires (near) real-time analytics revealing operational problems that can be acted upon immediately. Diagnostics based on historical information may help redesign processes, but do not solve current problems. Also, diagnostics should focus on problems that can still be influenced.
8. The challenges related to *organizational change* are numerous. Employees might resist change for a variety of reasons (job security, status, local optimization, reward mechanisms, etc.). Also, teams may revert to old ways of working when change is not implemented correctly. Even the best technology *will not fix broken organizations and processes* if not supported by proper change management.

As the last point illustrates, many problems cannot be solved with technology. However, creating *transparency will facilitate cultural change* when managed properly. Moreover, recent developments in Machine Learning (ML) and Artificial Intelligence (AI) also create new opportunities to overcome some of the challenges above. Also, more information is available in digital form, providing a better alignment between the physical world and the digital world. Consider, for example, the Internet of Things (IoT), connecting physical objects to the Internet and providing new sources of event data. In the remainder, we will focus on the opportunities provided by predictive and generative AI leveraging Object-Centric Event Data (OCED) [3].

Object-Centric Process Mining (OCPM)

Traditional process mining has focused on operational processes having a clear case notion. We refer to this as *case-centric process mining*. A process is seen as a network of activities executed for cases. Examples of cases are customer orders, patient treatments, student loans, and credit card applications. This aligns well with mainstream process modeling languages like Business Process Modeling Notation (BPMN). Events in this setting refer to an *activity* executed for a *case* at a specific *point in time*. However, as the field of process mining is maturing, it becomes evident that this is an oversimplification of reality. *Object-Centric Event Data* (OCED) generalizes the traditional notion of event data in several ways [3,5,10]. Each event may refer to any number of objects, and objects may be involved in any number of events. These are the so-called *Event-to-Object* (E2O) relations. Moreover, objects may be related through *Object-to-Object* (O2O) relations. Both relationships can be qualified (i.e., have a label describing the relation). For example, an E2O relation may have the label "is executed by" or "is responsible for" and an O2O relation may have the label "is part of" or "is related to". Also, both events and objects are typed, and for a given type, there may be standard attributes. As before, events have a timestamp. Object-attribute values may also have a timestamp to represent updates.

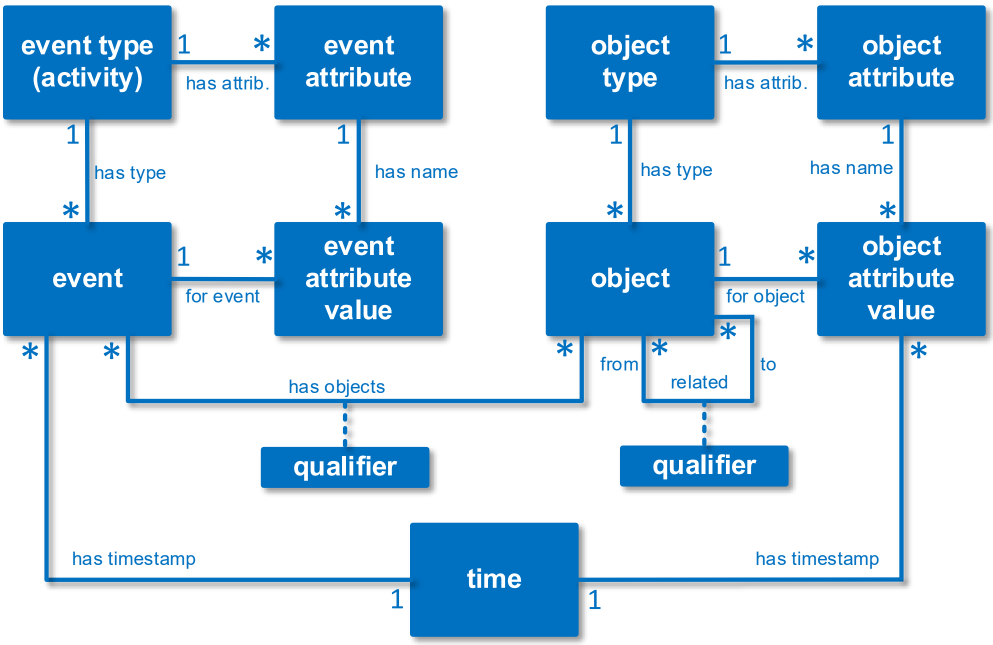

Figure 2: The OCEL 2.0 meta-model describing Object-Centric Event Data (OCED)

Figure 2 shows the *Object-Centric Event Log 2.0* (OCEL 2.0) meta-model [10]. OCEL 2.0 provides concrete storage and exchange formats for OCED, e.g., using JSON, XML, and relational and graph databases. As shown, events and objects are typed (see the two "has type" relations) and have attributes. The relation "has objects" connects events and objects. The relation is qualified and many-to-many to represent E2O relationships. The relation "related" connects objects to represent O2O relations.

Based on the OCEL 2.0 meta-model and earlier variants, a range of process mining techniques has been developed. However, compared to case-centric process mining, these are less mature, with many possibilities to improve them. The main *Object-Centric Process Mining* (OCPM) tasks are:

- **Object-Centric Process Discovery (OCPD)**: Automatically discovering object-centric process models from OCED. These models show the process flow of different object types. Examples are object-centric BPMN, object-centric Petri nets, object-centric process trees, and object-centric DFGs.

- **Object-Centric Conformance Checking (OCCC)**: Detecting and diagnosing commonalities and differences between observed OCED and modeled of discovered object-centric process models. This can be used to check compliance, taking into account different object types. Note that non-compliant behavior can be perceived to be compliant when considering objects in isolation.
- **Object-Centric Performance Analysis (OCPA)**: Analyzing the performance of processes involving multiple object types. As input, one can use OCED and/or object-centric process models annotated with performance information. The goal is to diagnose bottlenecks, compliance problems, and other performance or outcome-related problems. This also includes object-centric simulation.
- **Object-Centric Operational Support (OCOS)**: OCOS includes generating process predictions, process recommendations, and corrective actions. The goal is not to diagnose historical event data, but to create models that can be used operationally, e.g., predicting that an order will be delayed and actions are needed, or to foresee an emerging bottleneck. This often involves generating ML problems based on OCED and related process models. Having information on multiple object types increases the accuracy of such models.

The goal is not to detail the different OCPM techniques. What is evident from the above list is that many of the existing case-centric process mining techniques *need to be reinvented*.

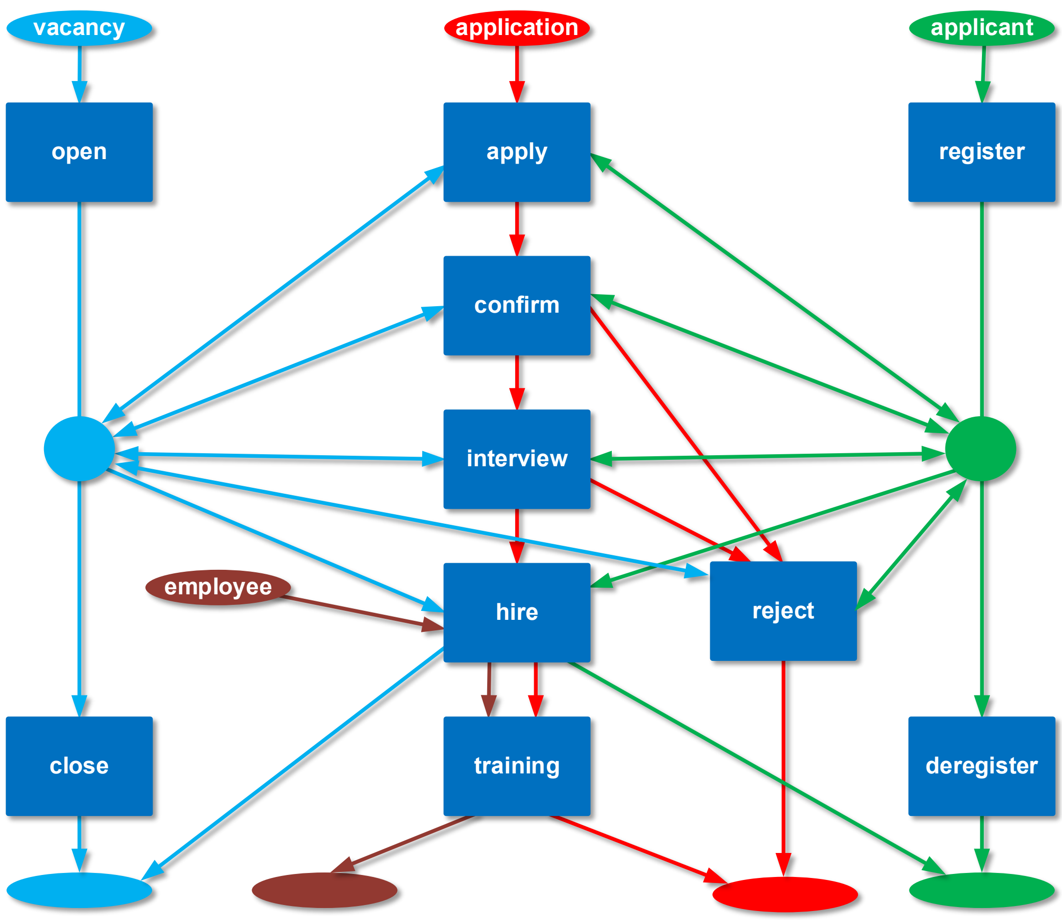

Figure 3: Simple artificial example to illustrate the basic OCPM concepts.

To illustrate the OCPM concepts, we use a simple example. Consider a hiring process involving the object types: *vacancy*, *applicant*, *application*, and *employee*. An *applicant* can *register* and *deregister*. Only registered applicants can *apply*. For a *vacancy*, we have the activities *open* and *close*. Applicants can only *apply* for an open vacancy. An event of type *apply* involves a *vacancy*, an *application*, and an *applicant*. As Figure 3 shows, there are also activities such as *confirm*, *interview*, *hire*, *training*, and *reject*. Events of type *confirm*, *interview*, *reject*, and *hire* all involve a *vacancy*, an *application*, and an *applicant*. Activities *hire* and *training* also involve an *employee*. Figure 3 sketches the interactions between these events and objects. For example, when hiring an *employee*, the *vacancy* is closed, and the *applicant* is deregistered. In this example, at most one object of a given type is involved in each event. However, it may also be the case that *multiple objects of the same type* are involved in a single event. For example, all remaining applicants are rejected in one step after hiring the first candidate for the position. Whereas case-centric process mining links each event to a single case, object-centric process mining poses no constraints on linking events and objects. Next to these Event-to-Object (E2O) relations, objects may be related through Object-to-Object (O2O) relations. These are not directly visible in a process model like in Figure 3, but are essential for filtering and selection.

Towards Domain-Specific Process Mining

The meta-model depicted in Figure 2 is generic and can be applied in logistics, production, finance, healthcare, etc. Although the meta-model is generic and somewhat abstract, concrete event data stored in such a format can be used directly due to a range of process mining techniques, e.g., discovering a process model. It is also possible to provide a standard set of performance indicators using generic concepts such as object type, event type, etc. For example, the average "lifetime" of objects of a given type (i.e., the difference between the last and the first event in which the object was involved). It is also possible to define specific performance indicators using languages such as PQL (Process Query Language) of Celonis. However, many organizations perform similar processes and it is not easy to create diagnostics for the different performance and compliance problems. Therefore, it does not make much sense to start from scratch when an organization starts with process mining. *Domain-specific process mining* aims to reuse preexisting knowledge and experiences and standardize the input. For example, it makes perfect sense to define the object and event types for a particular domain (e.g., sales, procurement, or human resources) in a unified manner. Let us return to the earlier rather simple example to explain this.

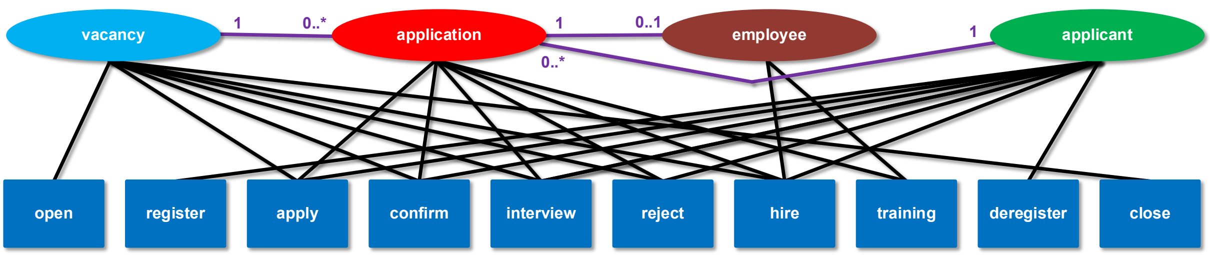

Figure 4: Object and event types and their relationships at the type level.

Figure 4 shows the four object types and ten event types (i.e., activities) used in the example. The figure also depicts the O2O relations in purple. For example, each application is related to precisely one vacancy and one applicant. For a vacancy, there may be multiple applications, and one applicant may have several applications (for different vacancies). In this example, there are many E2O relations cluttering the view a bit. Both O2O and E2O relations have qualifiers, but these are not shown. It is possible to specify the cardinalities of the E2O relations, but these are also not depicted. For example, for each vacancy, activity open is performed precisely once, and this is the only object involved. The hiring activity involves all four object types.

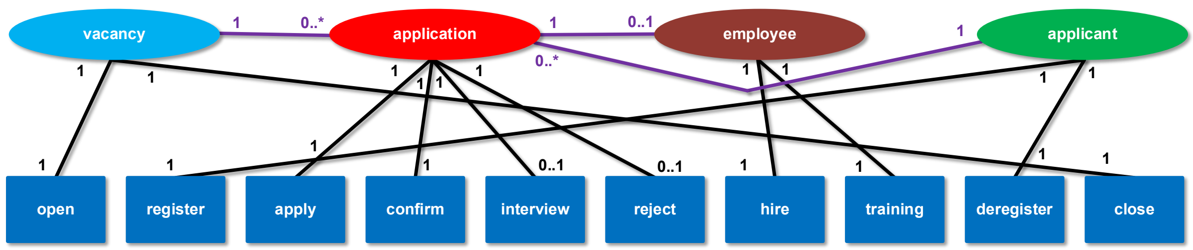

Figure 5: Simplification showing only the leading object type per event type.

Some of the relations in Figure 4 may be considered redundant. For example, if an event involves an application, it is possible to derive the corresponding applicant and vacancy using the O2O relations. Leaving out these derivable E2O relations yields a less cluttered view, as shown in Figure 5. Often it is desirable to pick one *leading object type* per event type such that the other objects involved can be derived from the leading object. Note in a process model like in Figure 3 one may still want to show the involvement of other objects next to the leading object. It is possible to define multiple criteria for explicitly including objects in an event (i.e., the E2O relations). For example, active participation of the object (e.g., an applicant being involved in an interview) and whether the object changes state (e.g., hiring an applicant for a vacancy is a relevant state change for both the vacancy and applicant) may be used as criteria.

The above discussion shows that one can model a specific domain by describing the object and event types, the E2O and O2O relations, and their cardinalities. There is no need to reinvent the wheel when dealing with processes such as Accounts Payable, Accounts Receivable, Inventory Management, Order Management, and Procurement. This is the reason Celonis and its partners provide over 300 so-called apps built on top of the Celonis process mining platform. These are still based on case-centric process mining and often assume a specific source system (e.g., SAP or Oracle). However, OCPM provides an opportunity to rethink these apps. The goal is to create a system-agnostic single source of truth based on a *domain-specific reference model*. Such a reference model consists of (at least) the following elements:

- A set of *object types* with a specification of the attributes of these object types.
- A set of *event types* (i.e., activities) with a specification of the attributes of these event types.
- A set of *Object-to-Object* (O2O) relations at the type level, including allowed cardinalities (one-to-one, one-to-many, etc.).
- A set of *Event-to-Object* (E2O) relations at the type level, including allowed cardinalities (e.g., a create order event refers to one sales order and at least one product and is executed only once for each sales order).
- A set of *performance indicators*, i.e., predefined functions computed over the events and objects specified above. Consider, for example, “On Time & In Full” (OTIF), which is the percentage of orders delivered on time and in full.
- A set of *normative process models*. Each process model refers to a subset of object types and event types. These models can be used for replaying predefined subsets of event data to show performance and compliance problems.

Note that a process is a “view” on a possibly larger collection of object and event types. In case-based process mining, a process and process model were defined by the case notion. In OCPM, there may be many possibly overlapping predefined views. Such *process lenses* provide the input for Machine Learning (ML) and Artificial Intelligence (AI).

Enabling Predictive AI

ML and AI are closely related fields, but they have distinct focuses. AI is a broad area of computer science aimed at creating machines capable of performing tasks that would typically require human intelligence. AI systems can be rule-based and deterministic, or they can learn and adapt over time. ML, on the other hand, is a subset of AI that focuses specifically on the development of algorithms and statistical models that enable computers to perform tasks without being explicitly programmed for this. The learning process involves analyzing patterns in the data and exploiting these. ML is the driving force behind many AI systems' ability to adapt and improve using data. Therefore, the terms are often used interchangeably. Generative AI focuses on creating new data or patterns, like ChatGPT, which generates human-like text. Predictive AI, on the other hand, analyzes existing data to predict future outcomes.

Predictive AI relies on historical data to learn patterns and relationships. The output is usually a specific prediction, a proposed decision (recommendation), or a probability score, indicating the likelihood of a future event or outcome. In most cases, only data specific to the problem are used. Many approaches can be described in abstract terms as learning a function $f \in X \to Y$ based on example inputs of the form $\{(x_1, y_1), (x_2, y_2), \dots (x_n, y_n)\}$ with goal of minimizing the error (e.g. minimize $\sum_i |y_i - f(x_i)|$). Often, the X value is composed of several input features (i.e., a vector of attributes). The Y value refers to the target feature one would like to predict or understand based on the X value. It is often impossible to directly apply such techniques to the data in source systems supporting processes. Often multiple systems are used and some of these systems may have tens of thousands of tables. However, OCPM provides the *process lenses* required to generate the input for learning a function $f \in X \to Y$.

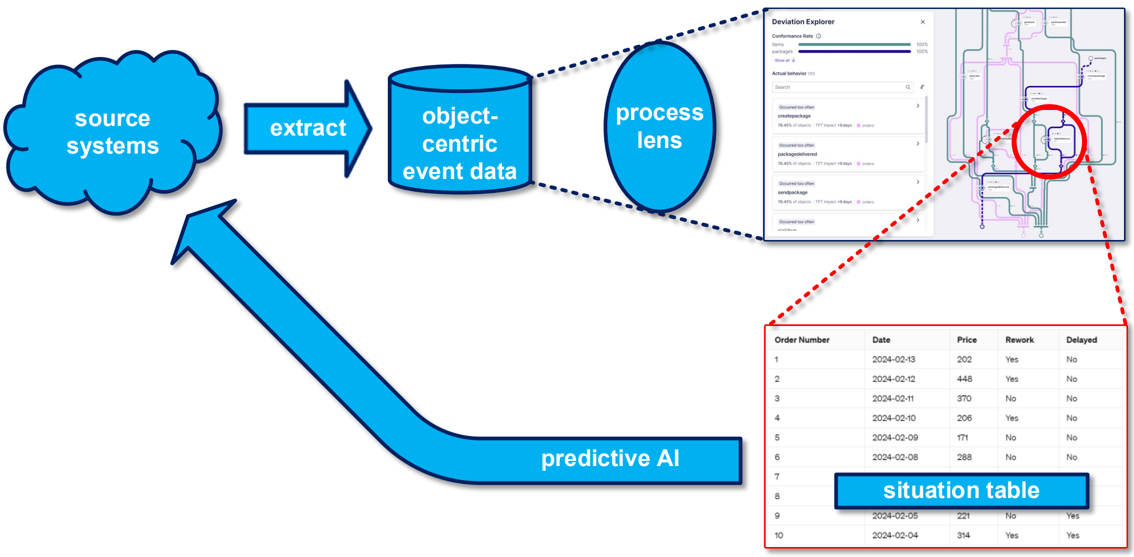

Figure 6: Relating process mining and predictive AI.

Figure 6 explains the interplay between OCPM and predictive AI. Data is extracted from the source systems, possibly guided by a domain-specific reference model. By picking subsets of object and event types, one can discover process models or check conformance with respect to a normative process model defined for that view. By replaying event data on such models (discovered or normative), one can identify performance and compliance problems. Given a problem of interest, we need to create a so-called *situation table*. For example, an organization would like to improve the "On Time & In Full" (OTIF) measure. Each row in the situation table corresponds to an order. The input features may include information on who worked on the order, what kind of products were ordered, routing information, etc. The target feature indicates whether the order was delivered on time and in full. Using mainstream ML techniques, it is possible to create a function $f \in X \rightarrow Y$ that predicts whether an order is likely to be late or incomplete. This can be used to uncover the root causes of the problem better or use the function to trigger actions if needed. Over the last 15 years, many process-related problems have been investigated using the approach depicted in Figure 6 [1,2,4,7]. However, only recently has the focus shifted from case-centric event data to object-centric event data [3,10]. This is a significant development because it will lower the effort needed to generate ML problems, and most process problems involve multiple types of objects. However, the core process mining techniques (e.g., process discovery and conformance checking) are very different from mainstream ML techniques like neural networks. Hence, OCPM and ML complement each other.

Enabling Generative AI

Generative AI (GenAI) focuses on creating new content or data that are similar to, but distinct from, the training data. GenAI can generate text, images, music, and more. Unlike predictive AI, the focus is not on a specific phenomenon, and huge amounts of data not specific to the problem are used. Large Language Models (LLM), like ChatGPT, try to generate new, original content that mimics the learned data. GenAI combines many very sophisticated ideas, but the best way to understand the mechanisms is to look at *n*-grams. An *n*-gram is a sequence of *n* words. For example, "I love to eat pizza in Italy" is a 7-gram. One can scan the internet to see how frequent *n*-grams are. Considering *n*-grams and *(n+1)*-grams, one can estimate the probability of the next word following a sequence of *n* words. This means that given an incomplete sentence, one can pick the word with the highest probability (or randomly pick a word based on the estimated probabilities). For example, the next word after seeing "I love to eat pizza in" is likely to be "Italy". This can be estimated by looking at all 7-grams starting with the first 6 words. By repeatedly picking the next word, the text grows and will seem surprisingly coherent. ChatGPT, Google Bard AI, Microsoft Copilot, Bing Chat, etc. are, of course, much more sophisticated but can be seen as ways of simply completing a sentence or text.

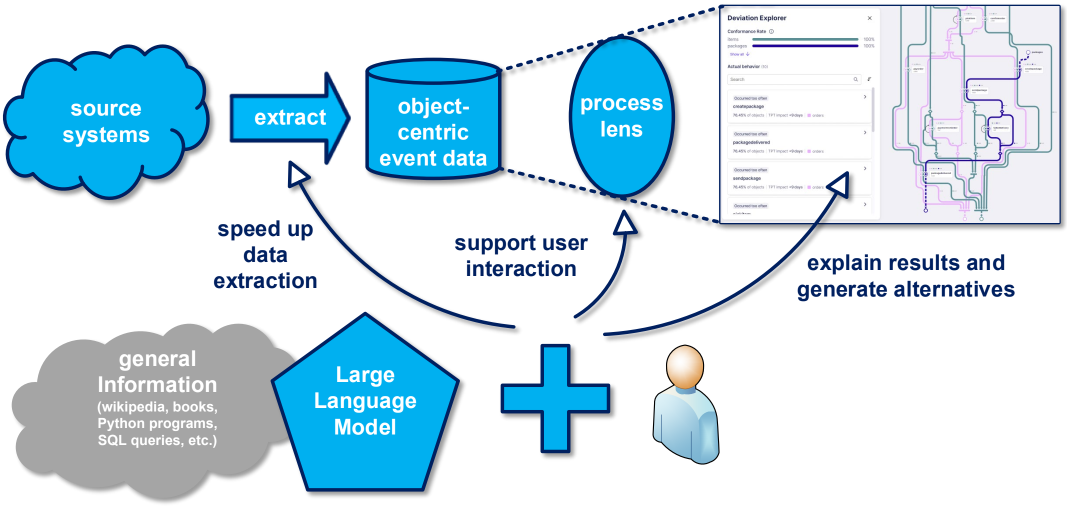

Figure 7: Using generative AI to lower the threshold to do process mining.

To use GenAI, one needs to create a prompt. There are many different prompt strategies, and prompt engineering is an active field of research [6]. The prompt may contain (1) actual data, (2) aggregated data, or just (3) metadata. Actual data are the actual events and objects. Aggregated data are summarizations of the event data. Examples are:

- The top *process variants* with their frequency and average duration. A variant may be the sequence <register, apply, confirm, apply, confirm, interview, reject, accept, training> of events for an applicant. The variant may have a frequency of 500 and an average duration of 2 months. This information could be provided in the prompt.
- The *directly-follows relations* with their frequency and average duration. For example, activity interview is directly followed by activity reject for 400 applicants and the average duration is two weeks.

Metadata only describe the structure, e.g., the database tables, the columns, and the key relationships. There is no information on individual events and objects, nor is there any information about frequencies and durations.

The prompt also needs to contain a question, e.g., Why are so many cases delayed? The output can be free text or in a specific format (e.g., SQL or PQL). There are also additional dimensions to describe prompting strategies, e.g., zero-shot, single-shot, and few-shot prompting. Zero-shot promoting means that no examples are given, relying on the LLM's general understanding. Single-shot promoting means that there is one example to guide the response. Few-shot prompting uses multiple examples. Iterative prompting means refining prompts based on previous LLM responses in an interactive way.

GenAI will not replace the core process mining techniques like process discovery and conformance checking, just like GenAI will not replace calculators or ILP solvers. However, GenAI will make it easier to interact with process mining tools. This is indicated in Figure 7. GenAI will make it simpler to extract event data (e.g., generate queries extracting event data). GenAI will make it easier to ask questions in natural language. Finally, it will help to explain process mining diagnostics. Although interactions become easier, the user always needs to verify that the answer is correct. Natural language is ambiguous, and GenAI suffers from so-called hallucinations. Organizations that have major data quality problems and a poor understanding of their processes should first get the basics right before using GenAI. Organizations need to crawl before they walk and walk before they run.

Outlook

Object-Centric Process Mining (OCPM) and Artificial Intelligence (AI) complement each other. Despite spectacular advances in AI, it is unrealistic that AI can answer questions about processes without taking a process-centric view on data scattered over tables in different IT systems. OCPM is a significant step forward compared to case-centric process mining. Both object and event data need to be stored uniformly, independent of the source systems. This uniformity eliminates the need for repeated data extraction whenever there is a change in perspective or a new case notion. Forcing complex intertwined processes into process models based on a singular case concept often results in misleading diagnostics. Such diagnostics only make sense for experts familiar with the data transformations applied. Moreover, OCPM enables the visualization and comprehension of interactions across different object types, emphasizing that performance and compliance issues cannot be understood when objects are considered in isolation.

A domain-specific reference model further structuring object-centric event data helps to create a coherent starting point for process mining and AI applications. Such a reference model predefines the different object and event types, attributes, O2O and E2O relations and their cardinalities, performance indicators, and a collection of normative process models. Using predictive AI, it is possible to diagnose performance and compliance problems. Here, OCPM helps identify the problems and provides the "process lenses" required to generate ML problems. Also Generative AI (GenAI) will benefit from this. On the one hand, the reference model helps to guide LLMs. On the other hand, GenAI makes it easier to interact with process mining tools. Therefore, we advocate a combination of AI and OCPM. *Predictive and generative AI complement process mining*. However, the core process mining capabilities are still needed. AI will *not* replace classical computation, e.g., calculators and ILP solvers. One does not want to guess what the sum of two numbers is. The same applies to process discovery and conformance checking. Moreover, calculators did not replace mathematicians, and AI will not replace domain experts. Therefore, process diagnostics need to be understandable by humans.

References

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

[2] W.M.P. van der Aalst. Hybrid Intelligence: To Automate or Not to Automate, That is the Question. International Journal of Information Systems and Project Management, 9(2):5-20, 2021.

[3] W.M.P. van der Aalst. Object-Centric Process Mining: Unraveling the Fabric of Real Processes, Mathematics, 11(2):2691, 2023.

[4] W.M.P. van der Aalst and J. Carmona, editors. Process Mining Handbook, volume 448 of Lecture Notes in Business Information Processing. Springer-Verlag, Berlin, 2022.

[5] W.M.P. van der Aalst and A. Berti. Discovering Object-Centric Petri Nets. Fundamenta Informaticae, 175(1-4):1-40, 2020.

[6] A. Berti, D. Schuster, W.M.P. van der Aalst. Abstractions, Scenarios, and Prompt Definitions for Process Mining with LLMs: A Case Study. In J. De Weerdt, L. Pufahl (eds) Business Process Management Workshops. BPM 2023. Lecture Notes in Business Information Processing, vol 492. Springer-Verlag, Berlin, 2023.

[7] M. Dumas, F. Fournier, L. Limonad, et al. AI-augmented Business Process Management Systems: A Research Manifesto. ACM Transactions on Management Information Systems, 14(1):1–19, 2023.

[8] IEEE Task Force on Process Mining. Process Mining Manifesto. In F. Daniel, K. Barkaoui, and S. Dustdar, editors, Business Process Management Workshops, volume 99 of Lecture Notes in Business Information Processing, pages 169-194. Springer-Verlag, Berlin, 2012.

[9] M. Kerremans, K. Iijima, A. Sachelarescu, N. Duffy, and D. Sugden. Magic Quadrant for Process Mining Tools, Gartner Research Note GG00774746, www.gartner.com, 2013.

[10] Process and Data Science Group. Object-Centric Event Log (OCEL 2.0) Website Providing Data Sets, Software, Etc. www.ocel-standard.org, 2023

[11] L. Reinkemeyer, editor, Process Mining in Action: Principles, Use Cases and Outlook. Springer-Verlag, Berlin, 2020.