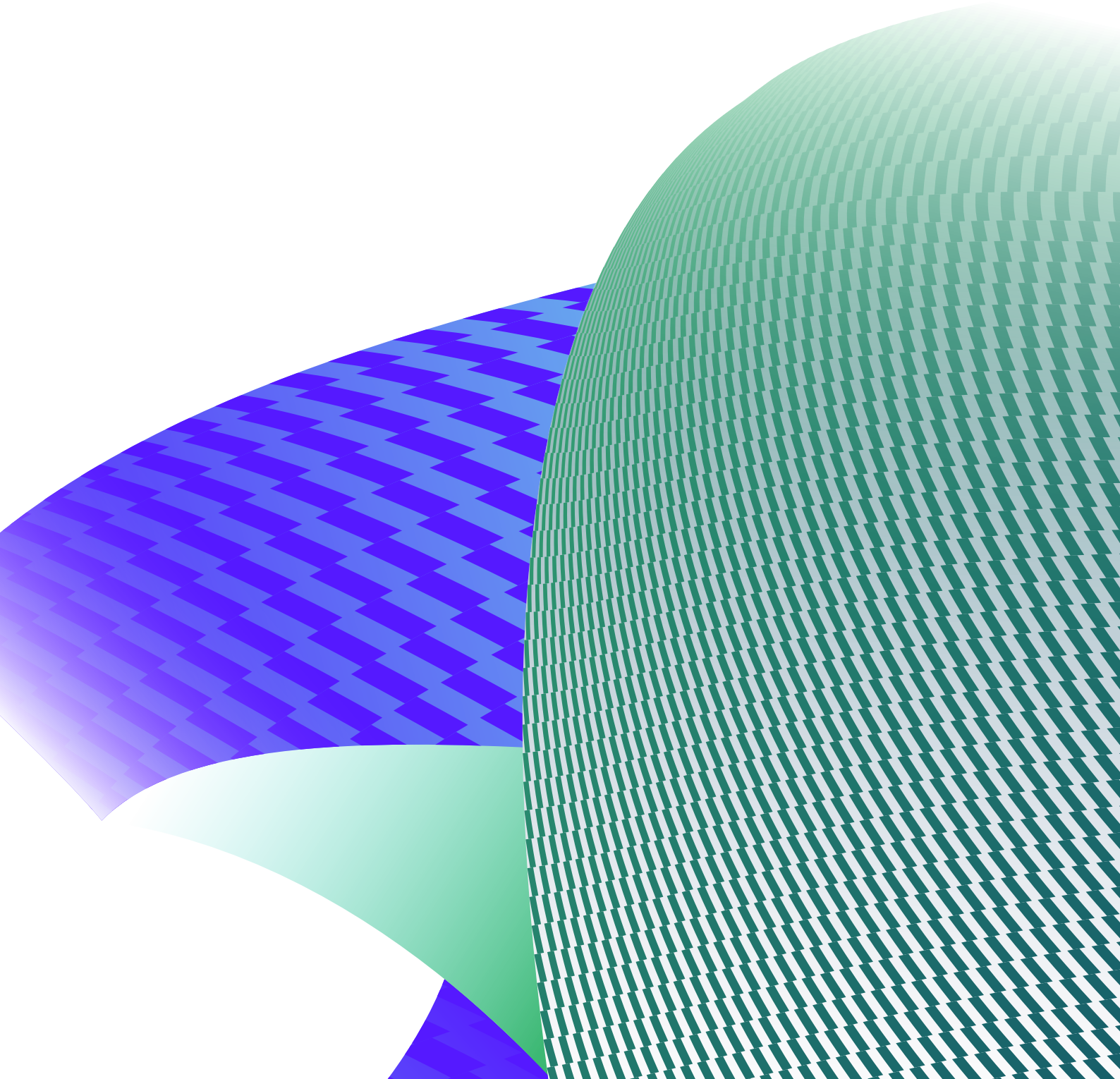
ABBYY

Different Approaches to Process Analytics

Timeline Analysis™ vs Schema Analysis



Timeline Analysis vs Schema Analysis

Table of contents

Introduction

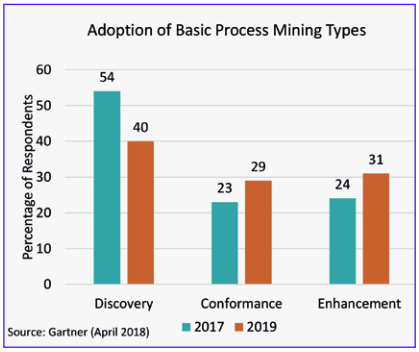
This document describes the key differences between two approaches to process analytics. The first approach, traditional process mining, features the schema approach. This method converts process data into a flowchart-like schema and then analyzes the flow of cases through that schema. The second approach, called Timeline Analysis, is based on the concept of a timeline–an unfiltered and unedited history of a single entity: a case, an opportunity, an order, a patient, or any other process.

This paper will demonstrate that the schema approach has many fundamental problems which severely affect its practical use in commercial application.

Process analytics for the modern business?

Process analytics is a specialized branch of business intelligence. It utilizes the same data as most other types of analytics utilities on the market today.

There have been several vendors introduced to this industry who are able to help companies with process mining and advanced analytics, and some include integrations with RPA and BPM or offer monitoring capabilities.
There are mainly two types of vendors: those that incorporate a schema-based approach, and those that incorporate the timeline approach. This whitepaper seeks to explore the differences between these two approaches.



However, it views data as the traces of business processes. It connects multiple records belonging to the same business entity (patient, online order, insurance claim, etc.) and allows for the visualization of the process data behind these records. This information is then delivered as a unique view of processes and can assist business users in better understanding processes with process insights that are not attainable with other types of BI.

Process analytics is closely tied with the concept of process mining. However, process mining is the foundation for real process analytics. With deeper analysis, it can create meaningful insights for an organization to more easily understand, manage, and improve business processes. Incorporating the use of artificial intelligence and machine learning allows the future of processes to be forecasted and can allow for alerts to be sent to avoid problems and prevent inefficiencies from ever occurring.

Gartner has identified process mining as a business trend for the past several years. It remains relevant due to companies being hesitant to adopt the technology, as well as the increasing pressures to use process mining paired with other trends. For example, with robotic process automation (RPA) becoming more widely adopted across different industries, companies are using process mining as the framework to identify opportunities for RPA and monitor success post implementation.

Process mining has been a hot topic for some years now; however, the needs of companies have changed. As the chart shows, it is projected that the reasons for adopting process mining will change. It is going beyond simply discovery and the needs are more aimed towards conformance and enhancement. Companies are seeking vendors who are able to cater to these changing needs.

How it works

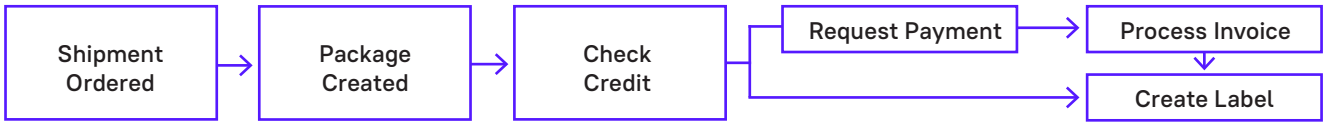
Let's take an example where a company sells something and must ship orders. As the company goes through the process–shipment being ordered, the package being created, etc.–data is being recorded.

Each record contains information including the instance, event, and time. This data is continuously collected throughout the business process. This data can be transformed to create a visualization of the process. This is where the data can be analyzed with process analytics.

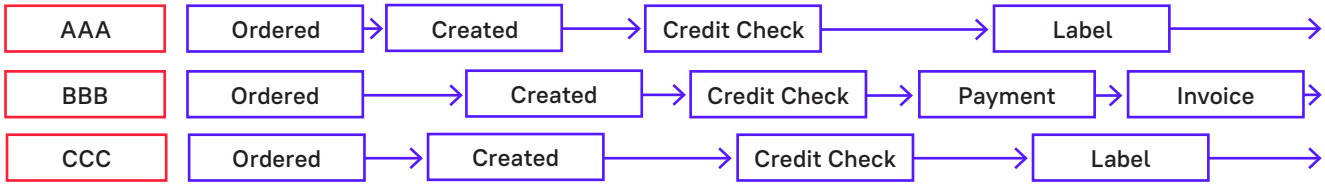
In both cases the data comes in the form of the records, each having at least the timestamp, instance identifier (unique identifier), and some description of what happened to this instance at this moment:

Time	Instance	Event	Some other attributes
02/13/2018 10:21 AM	AAA	Shipment Ordered	BOS-LAX
02/13/2018 11:40 AM	BBB	Shipment Ordered	NYC-SJC
02/13/2018 11:45 AM	AAA	Package Created	
02/13/2018 11:52 AM	AAA	Label Created	
02/13/2018 11:59 AM	BBB	Package Created	BOS-LAX

Here the similarity ends. The schema approach converts the records into a flowchart which is the closest fit for all records under analysis. Then the schema becomes the "backdrop" for various analyses and visualizations.



Contrary to that, the Timeline approach creates as many individual histories as there are entities:



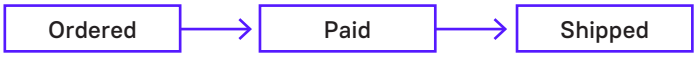
These timelines are analyzed by the tools just like BI analyzes the records in a table: compared, filtered, searched, aggregated, etc. In this approach, the schema is just one of the aggregated representations of the timelines. A user then can look at a process from any angle, including in a case management-like manner.

Problems and limitations and process schema approach

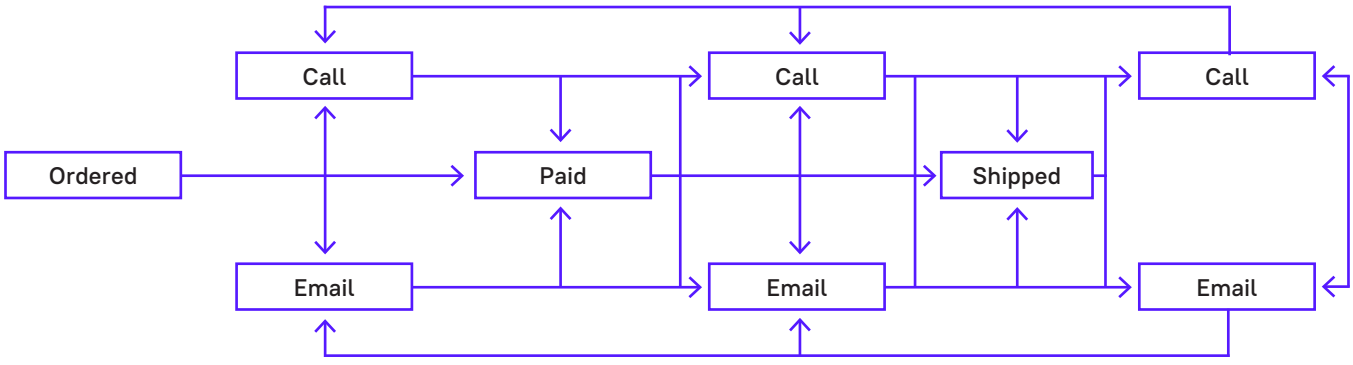
Generating the schema as the main goal of process analysis has been around for a very long time. It was initially considered a precursor for business process management (BPM): the common notion was "let's discover the existing schema and then feed it into the BPM tools for automation." This practice never materialized in commercial use since the complete reimplementation of processes is an unattainable task. So the same approach was applied to analytics. However, here it has several fundamental flaws.

1 Existence

The first problem with the schema is its existence itself. Not every process comes from the well-organized model that looks like a flow chart. Whether it's a case management or ad-hoc process or virtually any environment in which a decision for the next step is made on the previous step, the schema falls apart very quickly. That happens even in structured environments. Consider the simple process Ordered>Paid>Shipped. There's no trouble creating a schema for it as shown below:



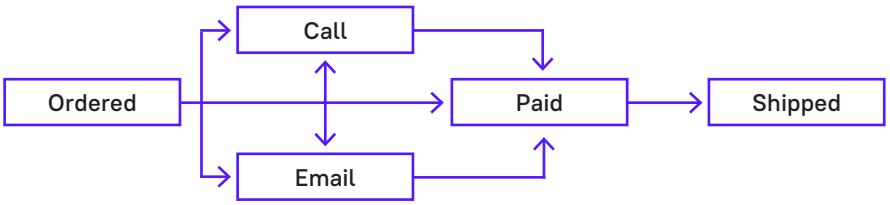
Now let's add just two additional actions: Call and Email. On every process step the worker could either progress to the next step or call or email the customer. The schema is converted into something that is difficult to comprehend and analyze for improvement opportunities:



And this is just for the five simple events. Imagine adding more.

2 Non-schema events

The second major issue of schema is that even for well-organized processes, some events do not fit into a schema nicely. Let's look again at the previous example. The actions Call and Email clearly do not have a good place in the schema. As a result, the schema-focused vendors do one of two things: they either remove these events completely and go back to the initial simplistic schema, or artificially put them in some place, just like this:

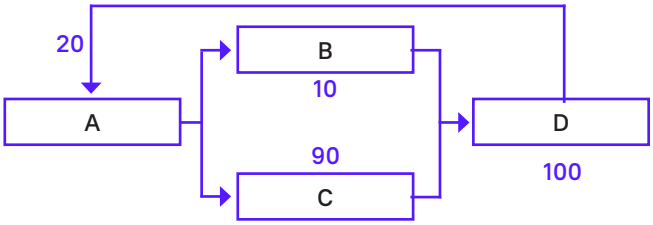


Neither solution makes much sense. If events are removed completely, the model loses important and sometimes critical information. Now we would never know that calls and emails constitute the largest activity in our business and the biggest contribution into the time and cost of the process. By placing the events into the "most likely" place in the schema, the method just confuses the user, because it's not the only place where such events occur.

3 Entity history

A huge drawback of the schema approach is that it doesn't show and therefore doesn't let users analyze the entire history of the entities, something that the timeline approach is fundamentally based on.

Look at the schema below:



The diagram shows that there were 20 transitions from D back to A, but here are just some of the questions requiring the knowledge of the histories of individual cases:

- Were there 20 cases which went back once, or 10 cases that had two setbacks each, or perhaps some other combination?
- If a case goes through the left branch B and then returns to A, would it go next time through B or through C?

The schema-based analysis is ill-suited to answer such questions because it fundamentally does not include the historical information of individual cases. The vast majority of process-related analytics vendors take this approach and limit the user in complete process understanding.

4 Compliance

Compliance is one of the most important reasons to invest in process analytics. Maintaining compliance means making sure things are happening as they should be, or how they're required to be.

In the schema approach, compliance analysis is based on the concept of the "happy path." It assumes there's one ideal path through the events and every deviation from it is a violation. This is huge simplification, rendering this approach useless in almost any real business environment.

The main reason for this is that compliance is not based on a path, but on the rules of the execution.

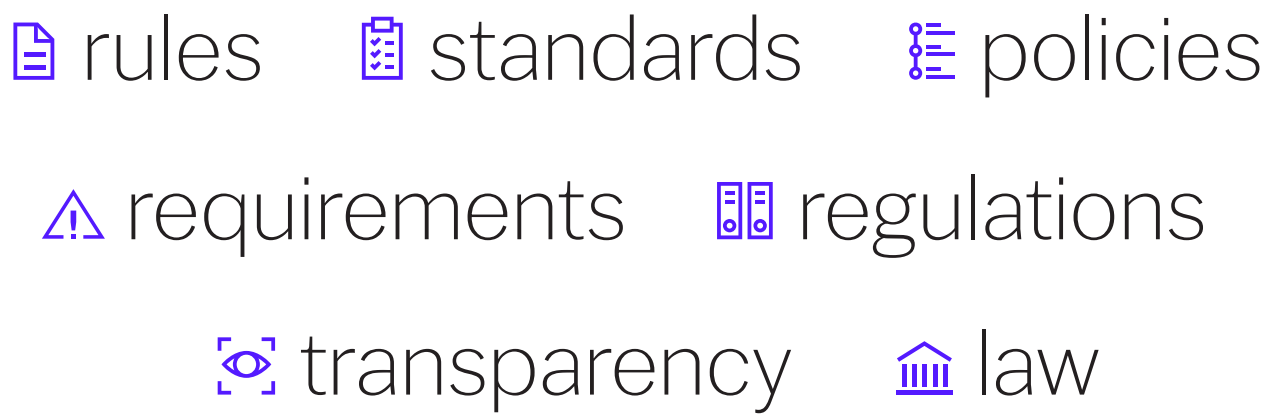
Instead of saying: you must do A B C D, the compliance rules say:

- If you do A, you should do B, C, and D.
- B and C could happen in any order, as long as they happen between A and D.
- D should happen no later than 3 days from A and 1 day from C.
- It is ok to perform D multiple times, but not C.

The happy path method fails to handle these requirements, therefore missing the mark in ensuring and analyzing process compliance for the enterprise.

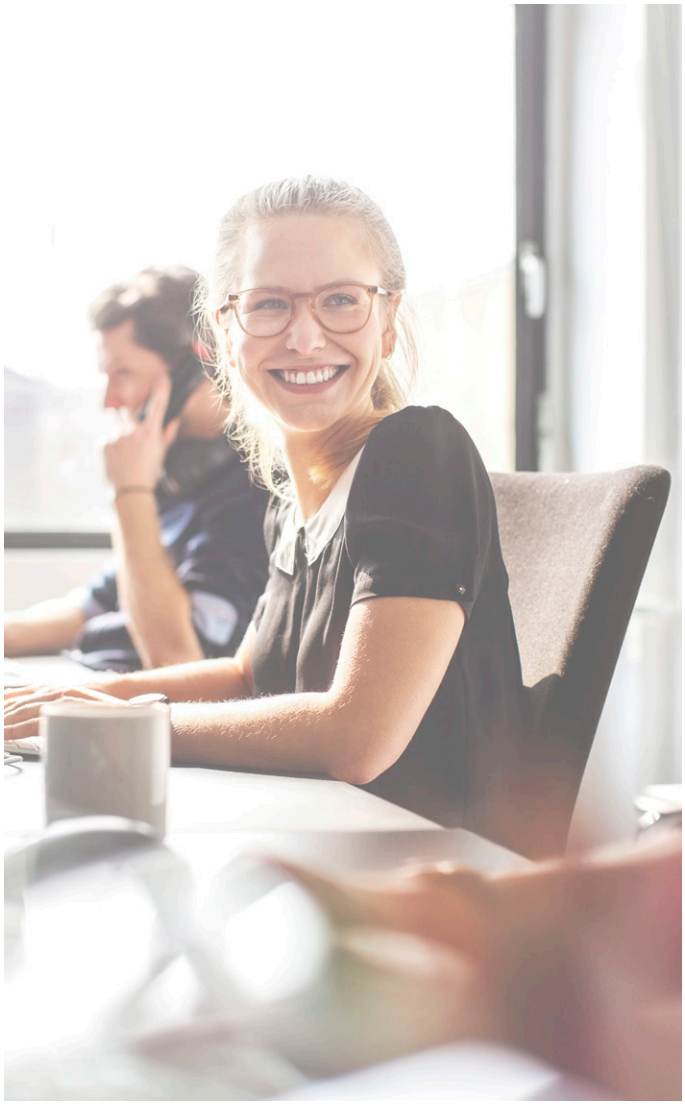
The second major problem with the happy path is that is doesn't describe the nature of the violations but rather declares them as deviations from the prescribed path. In a few cases it's ok; however, in the vast majority of businesses, violations are not created equal. For example, a minor delay in shipping is not the same as completely missing some prescribed step.

The combination of these two problems make the schema approach rather unusable for compliance analysis for the business user.



The timeline approach

As described above, the timeline approach makes the linear sequence of actions the focus of the analysis, essentially creating the histories of the objects. Nothing is filtered out or hidden. Then a plurality of highly specialized analysis methods is applied to those raw histories.



This approach was specifically designed to overcome the drawbacks of the schema-based approach.

This method has no trouble handling irregular and ad hoc processes since the history doesn't depend on the process nature. The non-schema events appear on the actual places in the histories and participate in all kinds of analysis. Entity history is a native representation of the timelines. Compliance is treated as a complex collection of rules which are applied to the histories timelines.

The timeline approach to process analytics provides 100 percent visibility of any process from end to end, even when different steps of the process are performed using multiple back-end systems. This enables the visualization and analysis of processes as executed, even when there is little to no integration between these systems.

This approach makes it easy to identify and quantify the impact of inefficient or broken process variations, determine their root causes and how to fix them, monitor and predict ongoing performance in real time to alert key staff for protocol deviations, and ensure positive outcomes.

The patented Timeline methodology

Compliance is one of the most important reasons to invest in process analytics. Maintaining compliance means making sure things are happening as they should be, or how they're required to be.

Existing analytical tools could be broken into categories of general business intelligence, process mining, and sequence analysis. General business intelligence tools analyze the relational data consisting of records and fields or rows and columns. Such tools are limited in that they do not provide visual interfaces and functionality to visualize and analyze the sequences of records.

Additionally, process mining tools reverse engineer the sequences of events into the process schema in the form of a business process model and notation (BPMN) or similar notations. While this task could be useful in cases where the business does have the underlying process schema, the task of discovering the common denominator for all sequences is different from understanding the variety and peculiarities of the actual sequences.

Further, existing sequence analysis tools are designed primarily for the field life sciences, specifically the analysis of the biological molecules such as DNA. Thus, existing sequence analysis tools do not incorporate the concept of time between events. Also, because of a very large number of elements in a single sequence, these tools rely on mathematical algorithms rather than a visual analysis to discover the patterns and deviations.

Raise your Process Intelligence with ABBYY Timeline

ABBYY Timeline is a comprehensive Process Intelligence platform offering over 30 prebuilt, best practice analysis tools to help business users understand business processes. The patented Timeline Analysis approach was built from years of experience in BI software, and frustrations with its limitations.

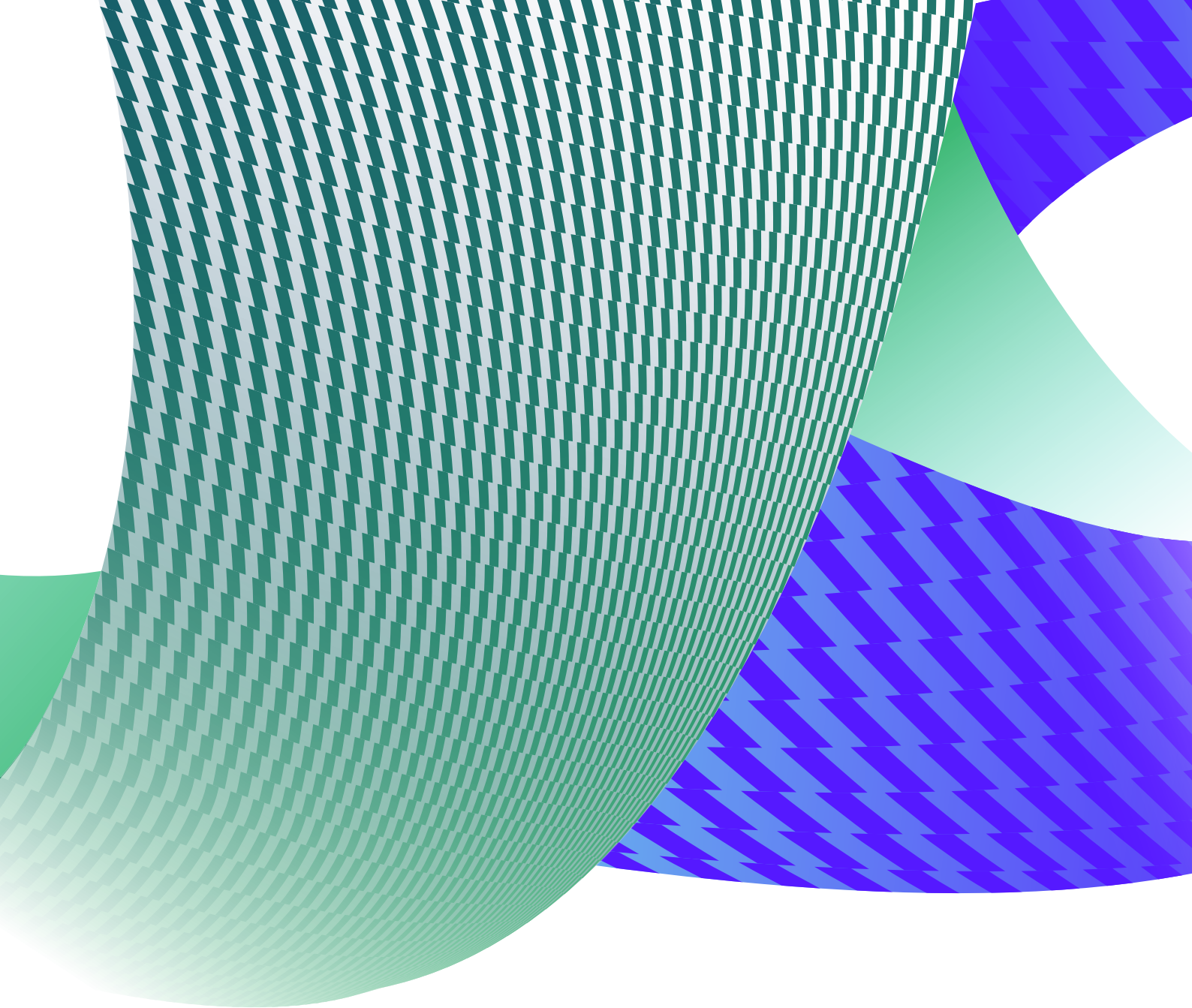
Analyze data for business-critical insights

- Improve customer journey
- Enable end-to-end transparency
- Reduce processing time
- Reduce compliance risks
- Ensure continuous improvement
- Reduce processing cost

Fast insights. Low effort. No risk.

- Fast Insights: See your processes like never before with sub-second response time.
- Low Effort: Get full support and guidance during your implementation/onboarding.
- No Risk: Try ABBYY Timeline for free.

For more Information, visit www.abbyy.com/timeline



ABBYY

WWW.ABBYY.COM