

Introduction to Process Mining

Dr.Raouf Khayami

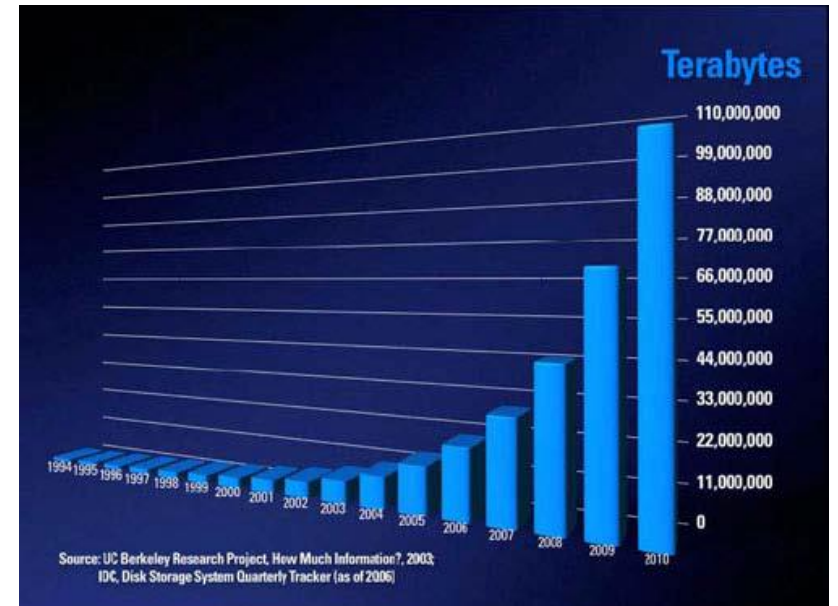
Fatemeh Davalloo

Index

- Introduction
- Process Mining
- Event Logs
- Process Discovery
- Prom

Introduction

- ▶ The growth of a digital universe that is well-aligned with process in organizations makes it possible to record and analyze events.
- ▶ Process mining provides new ways to utilize the abundance of data in enterprises.
- ▶ Process mining techniques have matured significantly over the past few years and as a result, management trends related to process improvement and compliance can now benefit from process mining.



What is a Business Process?

- ▶ Business processes are what companies do whenever they deliver a service or a product to customers.
- ▶ The way processes are designed and performed affects both “*the quality of service*” that customers perceive and the *efficiency* with which services are delivered.

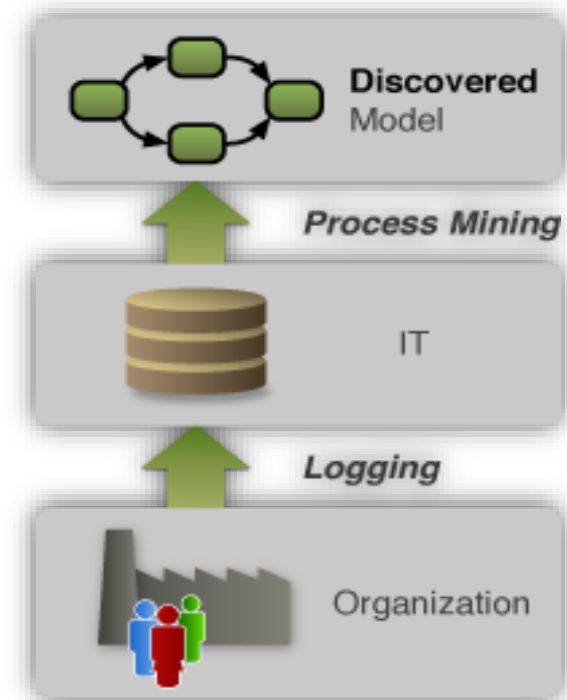
What is a business process?

- ▶ Business processes encompasses a number of events and activities, and involves a number of actors.
- ▶ Events correspond to things that happen atomically.



What is Process Mining?

- ▶ Today every organization have its own information system to store various kinds of data about their tasks and services.
- ▶ Process mining aims to discover, monitor and improve real processes by extracting knowledge from event logs readily available in today's information systems.
- ▶ Real process means the process that is actually taking place in an organization.



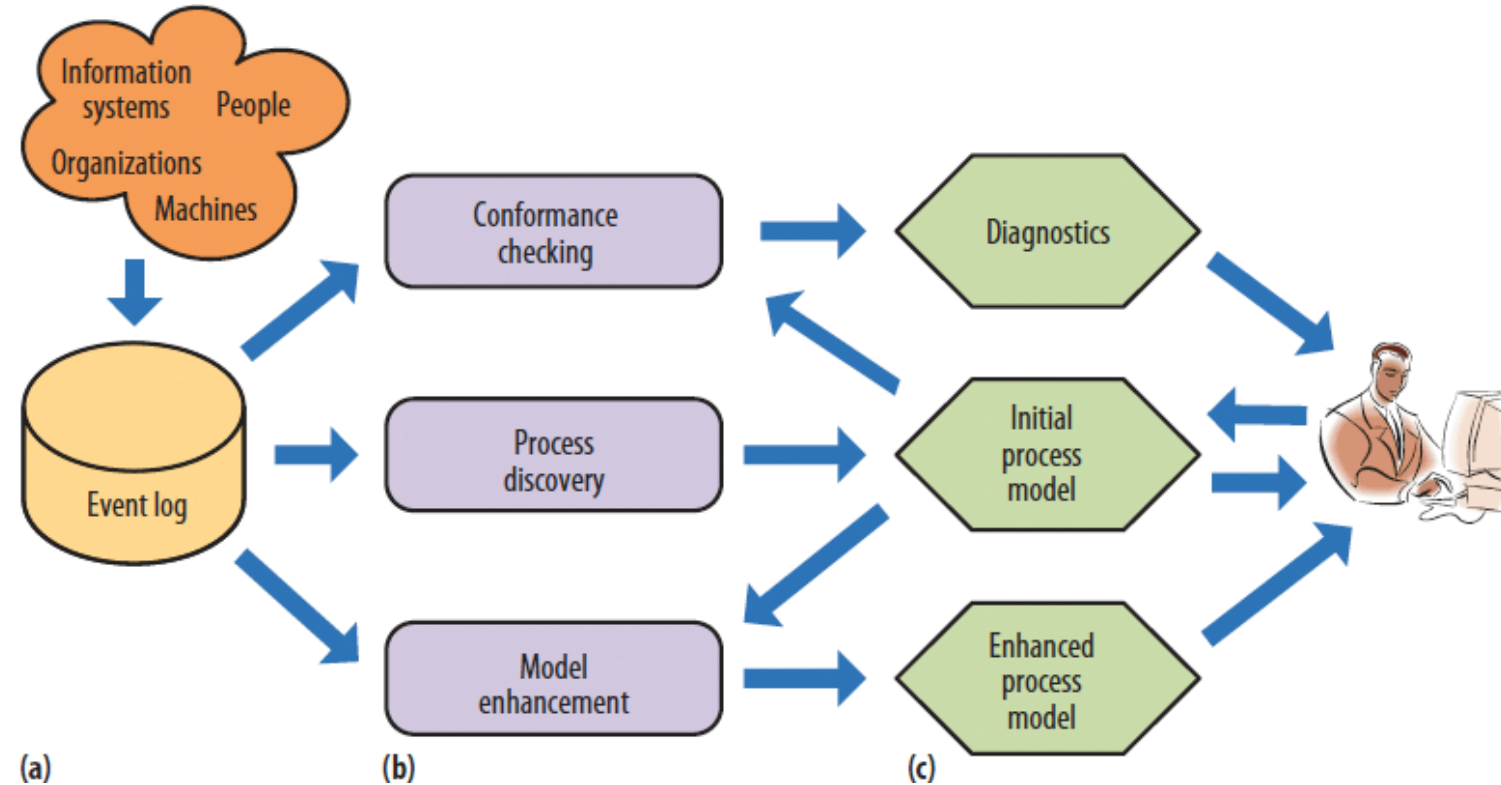
What is Process Mining?

- ▶ All of the process mining techniques assume that it is possible to sequentially record events in a way, that each event refers to an activity and is related to a particular case.
- ▶ These events and their detailed information make an event log.

case id	event id	properties				
		timestamp	activity	resource	cost	...
1	35654423	30-12-2010:11.02	register request	Pete	50	...
	35654424	31-12-2010:10.06	examine thoroughly	Sue	400	...
	35654425	05-01-2011:15.12	check ticket	Mike	100	...
	35654426	06-01-2011:11.18	decide	Sara	200	...
	35654427	07-01-2011:14.24	reject request	Pete	200	...
2	35654483	30-12-2010:11.32	register request	Mike	50	...
	35654485	30-12-2010:12.12	check ticket	Mike	100	...
	35654487	30-12-2010:14.16	examine casually	Pete	400	...
	35654488	05-01-2011:11.22	decide	Sara	200	...
	35654489	08-01-2011:12.05	pay compensation	Ellen	200	...
3	35654521	30-12-2010:14.32	register request	Pete	50	...
	35654522	30-12-2010:15.06	examine casually	Mike	400	...
	35654524	30-12-2010:16.34	check ticket	Ellen	100	...
	35654525	06-01-2011:09.18	decide	Sara	200	...
	35654526	06-01-2011:12.18	reinitiate request	Sara	200	...
	35654527	06-01-2011:13.06	examine thoroughly	Sean	400	...
	35654530	08-01-2011:11.43	check ticket	Pete	100	...
	35654531	09-01-2011:09.55	decide	Sara	200	...
35654533	15-01-2011:10.45	pay compensation	Ellen	200	...	
4	35654641	06-01-2011:15.02	register request	Pete	50	...
	35654643	07-01-2011:12.06	check ticket	Mike	100	...
	35654644	08-01-2011:14.43	examine thoroughly	Sean	400	...
	35654645	09-01-2011:12.02	decide	Sara	200	...
	35654647	12-01-2011:15.44	reject request	Ellen	200	...

Different Types of Process Mining

- ▶ Discovery
- ▶ Conformance
- ▶ Enhancement



Discovery

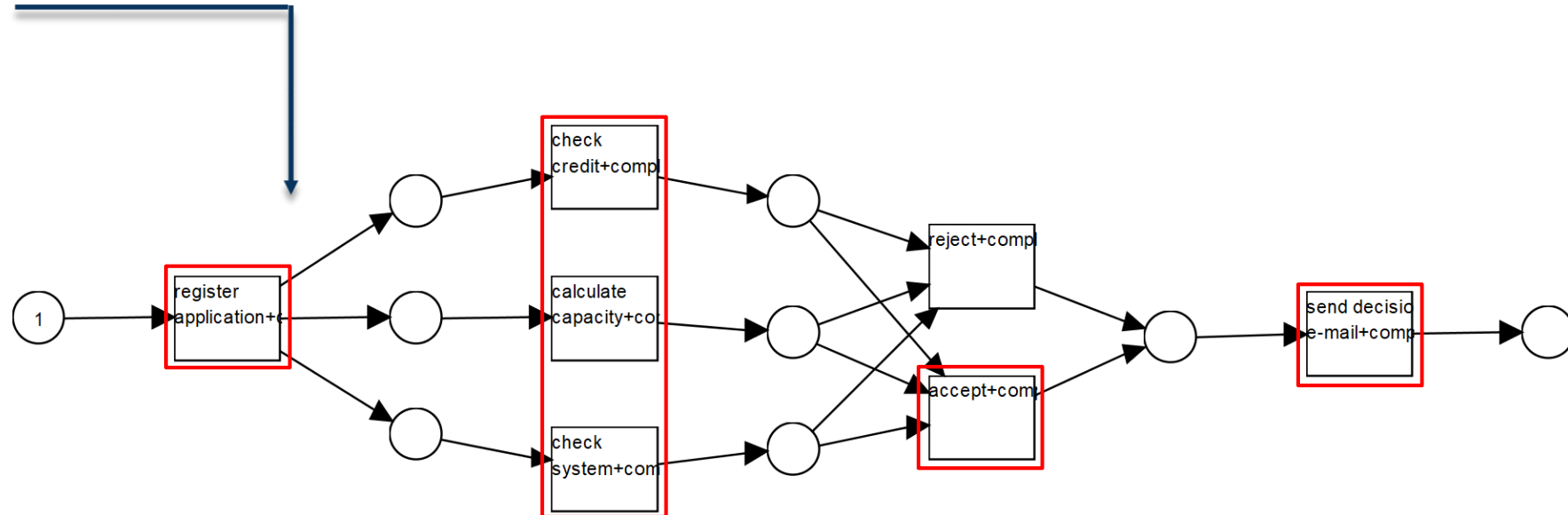
- ▶ Deriving information from some event log without using an a priori model.
- ▶ Based on an event log various types of models may be discovered.
 - Process model
 - Business rules
 - Organizational model



Discovery

▶ Example of a *process model* derived from an artificial loan event log.

Case ID	Properties event	Timestamp
0	register application	4/16/2013 10:08
0	check credit	4/16/2013 10:16
0	calculate capacity	4/16/2013 10:16
0	check system	4/16/2013 10:20
0	accept	4/16/2013 10:21
0	send decision e-mail	4/16/2013 10:26
1	register application	4/16/2013 10:10
1	check credit	4/16/2013 10:16
1	calculate capacity	4/16/2013 10:16
1	check system	4/16/2013 10:20
1	accept	4/16/2013 10:24
1	send decision e-mail	4/16/2013 10:29
2	register application	4/16/2013 10:15
2	check credit	4/16/2013 10:22
2	calculate capacity	4/16/2013 10:23
2	check system	4/16/2013 10:29
2	accept	4/16/2013 10:37
2	send decision e-mail	4/16/2013 10:44



Conformance Checking

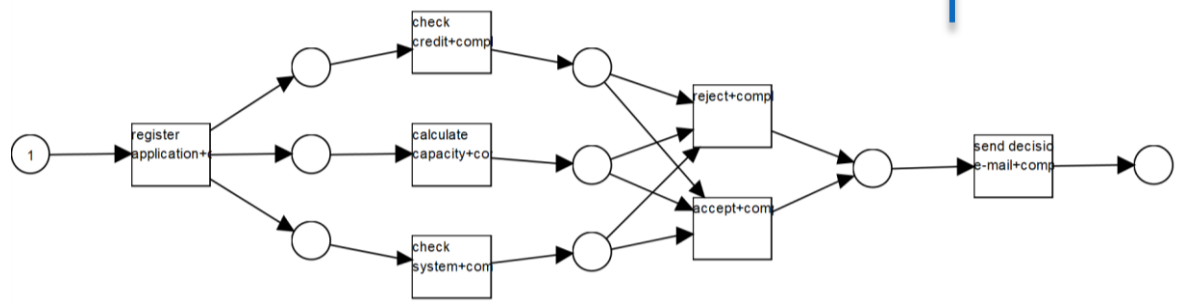
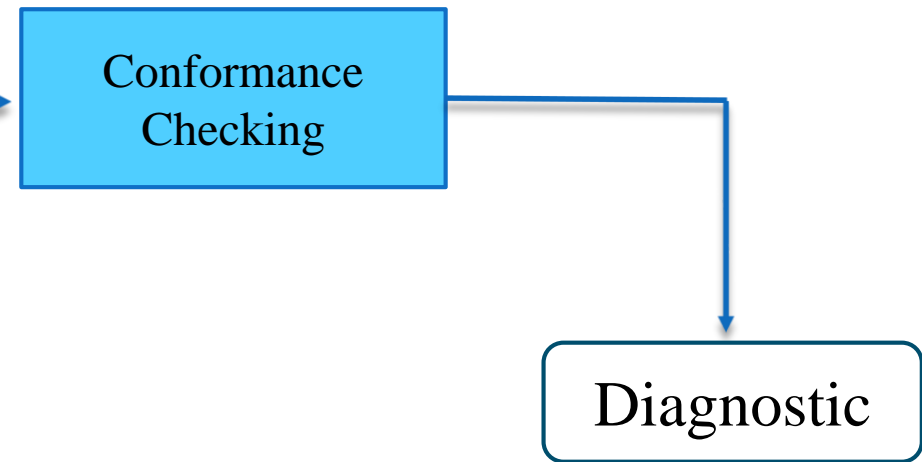
- ▶ An existing process model is compared with an event log of the same process.
- ▶ Conformance checking can be used to make sure reality, as recorded in the log, conforms to the model and vice versa.



Conformance Checking cont.

- ▶ A prerequisite of conformance analysis is that the tasks in *process model* must be *associated* with the *logged events*.

Case ID	Properties event	Timestamp
0	register application	4/16/2013 10:08
0	check credit	4/16/2013 10:16
0	calculate capacity	4/16/2013 10:16
0	check system	4/16/2013 10:20
0	accept	4/16/2013 10:21
0	send decision e-mail	4/16/2013 10:26
1	register application	4/16/2013 10:10
1	check credit	4/16/2013 10:16
1	calculate capacity	4/16/2013 10:16
1	check system	4/16/2013 10:20



Conformance Checking

- ▶ Conformance checking demands for two different types of metrics:
 - **Fitness:** the extent to which the logged traces can be associated with valid execution paths specified by the process model
 - **Appropriateness:** the degree of accuracy in which the process model describes the observed behavior, combined with the degree of clarity in which it is represented.

Enhancement

- ▶ Improve or extend an existing process model using information about the actual process recorded in some event log.
- ▶ E.g. by using timestamps in the event log one can extend the model to show bottlenecks, service levels, throughout time, frequencies.



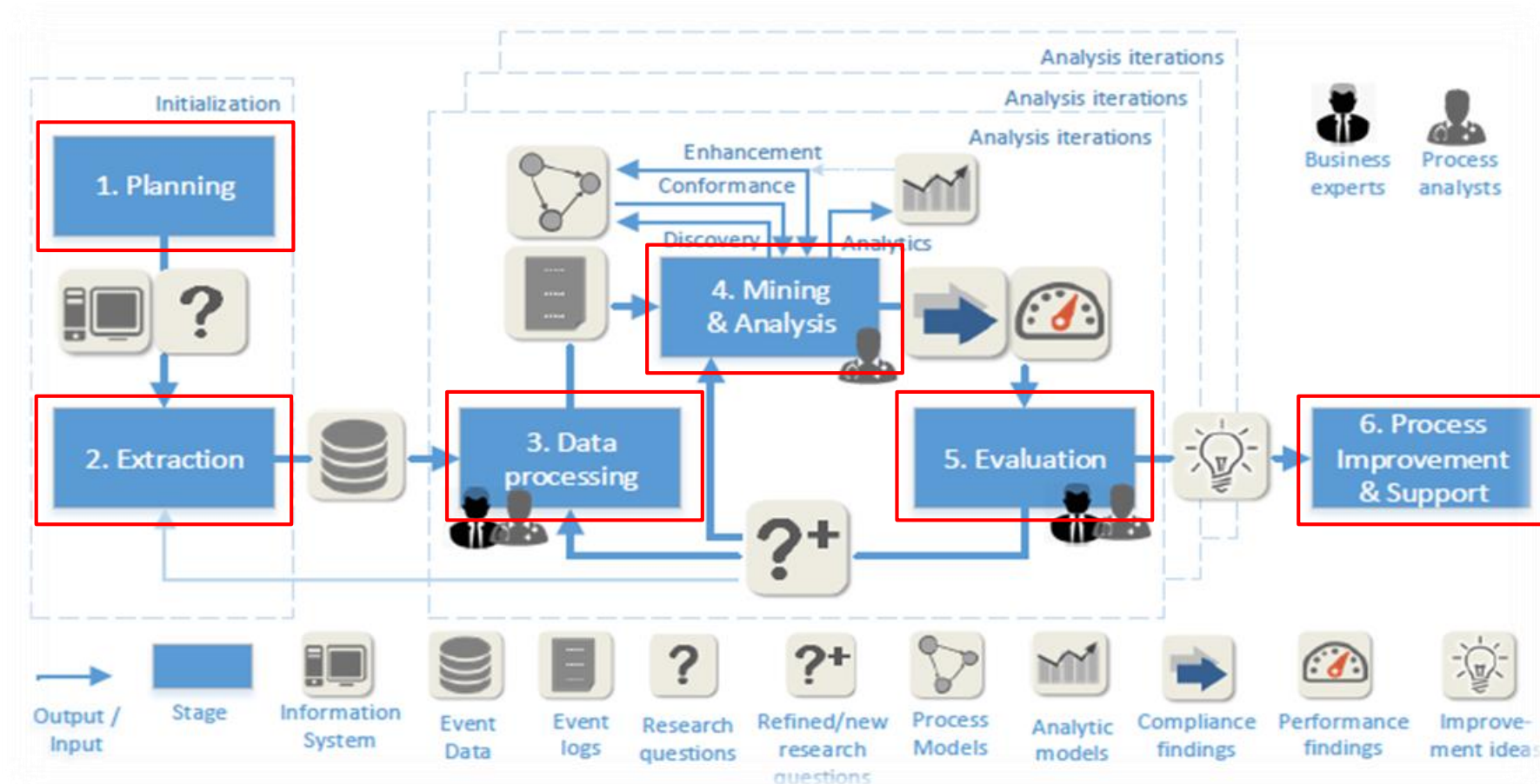
Process Mining Perspectives

- ▶ Process mining can also be categorized into four different perspectives

Control-flow	Focuses on the control flow, i.e. the ordering of activities.
Organizational	Focuses on information about resources hidden in the log, i.e. which actors are involved and how are they related.
Case	Focuses on properties of cases.
Time	Concerned with the timing and frequency of events.

PM2-a Process Mining Project Methodology

- ▶ PM2 is a methodology, designed to guide the execution of process mining projects.



Index

- ✓ Introduction
- ✓ Process Mining
 - ✓ Definition
 - ✓ Types
 - ✓ Perspectives
 - ✓ Methodology
- Event Logs
 - Concept
 - Formats
- Process Discovery
- Prom

Event Logs: starting point of process mining

▶ *Events* are recorded *sequentially*, each event refers to an *activity* and is related to a particular *case*.

- ▶ **Activity:** specific step in some process
- ▶ **Resource:** the person who executed the event
- ▶ **Case (Trace):** a process instance

case id	event id	properties				
		timestamp	activity	resource	cost	...
1	35654423	30-12-2010:11.02	register request	Pete	50	...
	35654424	31-12-2010:10.06	examine thoroughly	Sue	400	...
	35654425	05-01-2011:15.12	check ticket	Mike	100	...
	35654426	06-01-2011:11.18	decide	Sara	200	...
	35654427	07-01-2011:14.24	reject request	Pete	200	...
2	35654483	30-12-2010:11.32	register request	Mike	50	...
	35654485	30-12-2010:12.12	check ticket	Mike	100	...
	35654487	30-12-2010:14.16	examine casually	Pete	400	...
	35654488	05-01-2011:11.22	decide	Sara	200	...
	35654489	08-01-2011:12.05	pay compensation	Ellen	200	...
3	35654521	30-12-2010:14.32	register request	Pete	50	...
	35654522	30-12-2010:15.06	examine casually	Mike	400	...
	35654524	30-12-2010:16.34	check ticket	Ellen	100	...
	35654525	06-01-2011:09.18	decide	Sara	200	...
	35654526	06-01-2011:12.18	reinitiate request	Sara	200	...
	35654527	06-01-2011:13.06	examine thoroughly	Sean	400	...
	35654530	08-01-2011:11.43	check ticket	Pete	100	...
4	35654531	09-01-2011:09.55	decide	Sara	200	...
	35654533	15-01-2011:10.45	pay compensation	Ellen	200	...
	35654641	06-01-2011:15.02	register request	Pete	50	...
4	35654643	07-01-2011:12.06	check ticket	Mike	100	...
	35654644	08-01-2011:14.43	examine thoroughly	Sean	400	...
	35654645	09-01-2011:12.02	decide	Sara	200	...
	35654647	12-01-2011:15.44	reject request	Ellen	200	...
5	35654711	06-01-2011:09.02	register request	Ellen	50	...
	35654712	07-01-2011:10.16	examine casually	Mike	400	...
	35654714	08-01-2011:11.22	check ticket	Pete	100	...
	35654715	10-01-2011:13.28	decide	Sara	200	...
	35654716	11-01-2011:16.18	reinitiate request	Sara	200	...
	35654718	14-01-2011:14.33	check ticket	Ellen	100	...
	35654719	16-01-2011:15.50	examine casually	Mike	400	...
	35654720	19-01-2011:11.18	decide	Sara	200	...
	35654721	20-01-2011:12.48	reinitiate request	Sara	200	...
	35654722	21-01-2011:09.06	examine casually	Sue	400	...
	35654724	21-01-2011:11.34	check ticket	Pete	100	...
	35654725	23-01-2011:13.12	decide	Sara	200	...
35654726	24-01-2011:14.56	reject request	Mike	200	...	

Event Logs: Standard Formats

- ▶ Most of the information systems have their own format to record data, hence the need to create standard formats for event logs.

- ▶ MXML

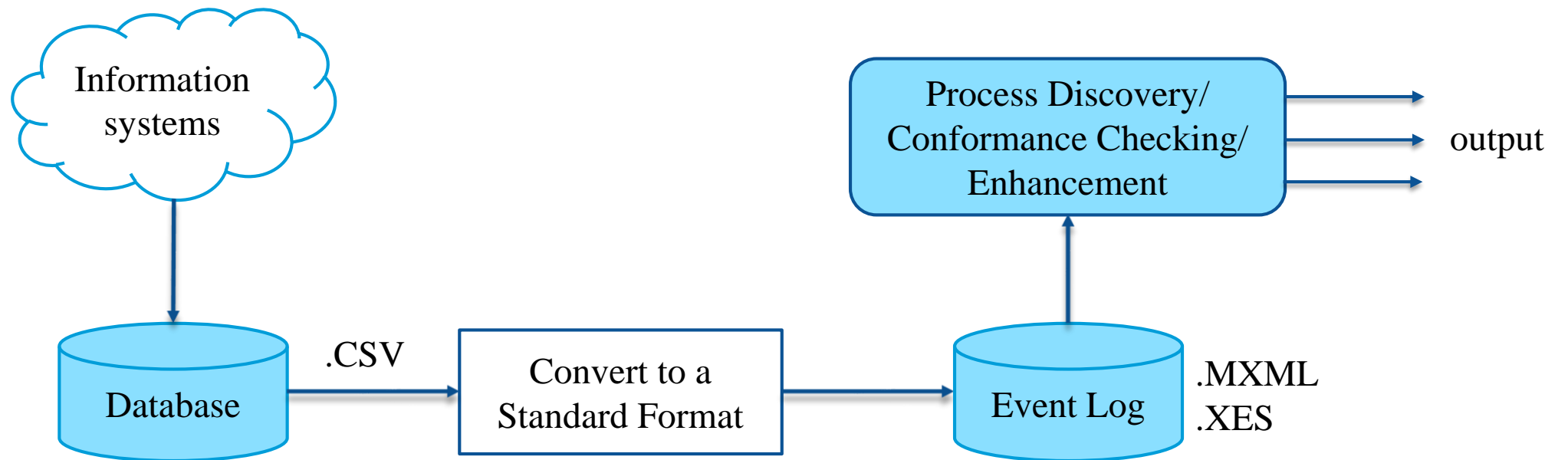
```
<ProcessInstance id="Order 1" description="instance with Order 1">
  <Data>
    <Attribute name="totalValue">2142.38</Attribute>
  </Data>
  <AuditTrailEntry>
    <WorkflowModelElement>Create</WorkflowModelElement>
    <EventType>complete</EventType>
    <originator>Wil</originator>
    <timestamp>2009-01-03T15:30:00.000+01:00</timestamp>
    <Data>
      <Attribute name="currentValue">2142.38</Attribute>
      <Attribute name="requestedBy">Eric</Attribute>
      <Attribute name="supplier">Fluxi Inc.</Attribute>
      <Attribute name="expectedDelivery">2009-01-12T12:00:00.000+01:00</Attribute>
    </Data>
  </AuditTrailEntry>
</ProcessInstance>
```

- ▶ XES

```
<trace>
  <string key="concept:name" value="Order 1"/>
  <float key="order:totalValue" value="2142.38"/>
  <event>
    <string key="concept:name" value="Create"/>
    <string key="lifecycle:transition" value="Complete"/>
    <string key="org:resource" value="Wil"/>
    <date key="time:timestamp" value="2009-01-03T15:30:00.000+01:00"/>
    <float key="order:currentValue" value="2142.38"/>
    <string key="details" value="Order creation details">
      <string key="requestedBy" value="Eric"/>
      <string key="supplier" value="Fluxi Inc."/>
      <date key="expectedDelivery" value="2009-01-12T12:00:00.000+01:00"/>
    </string>
  </event>
</trace>
```

Event Logs: Standard Formats cont.

- ▶ In order to use non-standard event logs created by some information system, first it should be converted to a standard one.



Event Logs: Challenges and Problems

- ▶ Uncompleted events
- ▶ No timestamp or sequential order
- ▶ Duplicated events
- ▶ Invisible Tasks
- ▶ etc.

Index

- ✓ Introduction
- ✓ Process Mining
 - ✓ Definition
 - ✓ Types
 - ✓ Perspectives
 - ✓ Methodology
- ✓ Event Logs
 - ✓ Concept
 - ✓ Formats
- Process Discovery
- Prom

Process Discovery

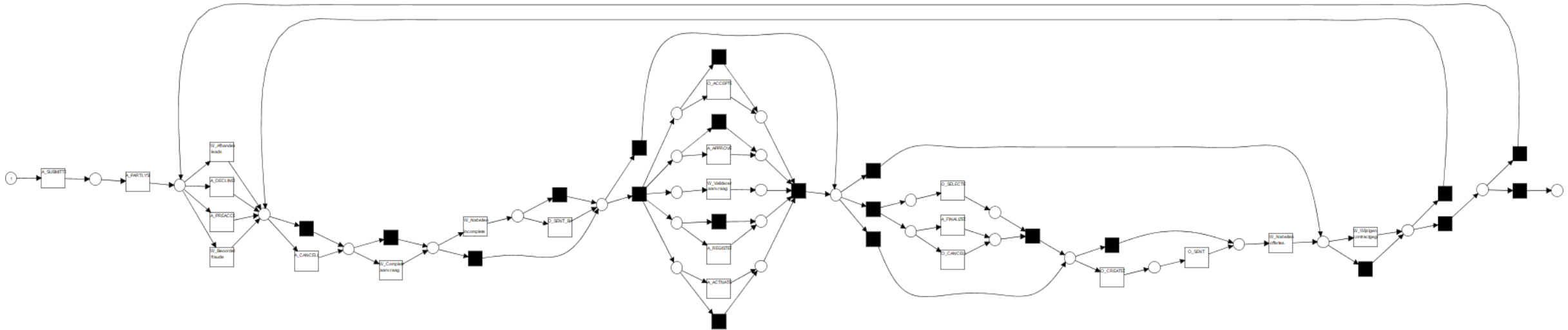
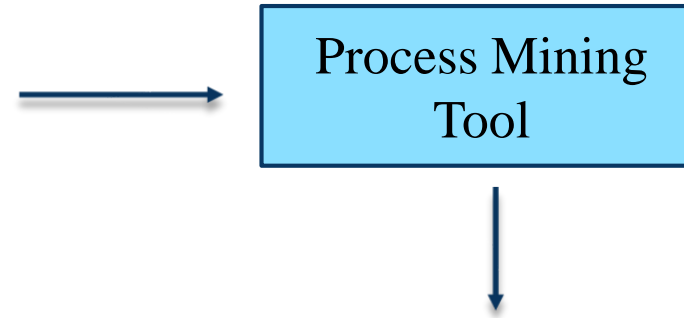
- ▶ In most cases the required data to create a standard event log is scattered through multiple tables or databases.
- ▶ To prepare an event log:
 - Have plan and research questions
 - Study the information system, databases and tables
 - Choose required attributes based on your research questions
 - Find the mutual attributes to connect tables
 - Identify cases, events and their properties

Process Discovery

- ▶ Real event log on handling loan applications for small consumer loans.
- ▶ Tables:
 - Personal Profile
 - Submitted Requests
 - Handling Requests
 - etc.

Process Discovery: example

case	event	startTime	completeTime	REG_DATE	AMOUNT_REQ
173688	SUBMITTED	38:44.5	38:44.5	38:44.5	20000
173688	PARTLYSUBMITTED	38:44.9	38:44.9	38:44.5	20000
173688	PREACCEPTED	39:37.9	39:37.9	38:44.5	20000
173688	PREACCEPTED	39:37.9	39:37.9	38:44.5	20000
173688	ACCEPTED	42:43.3	42:43.3	38:44.5	20000
173688	SELECTED	45:09.2	45:09.2	38:44.5	20000
173688	FINALIZED	45:09.2	45:09.2	38:44.5	20000
173688	CREATED	45:11.2	45:11.2	38:44.5	20000
173688	SENT	45:11.4	45:11.4	38:44.5	20000
173688	SENT	45:11.4	45:11.4	38:44.5	20000



Process Mining Application

- ▶ Several process mining tools have been developed over the past few years, here are some examples

Process Mining Tool	Vendor	Website
<i>ProM</i>	<i>Open source</i>	www.promtools.org
ProMLite	Open source	www.promtools.org
RapidProM	Open source	www.rapidprom.org
Celonis Process Mining	Celonis GmbH	www.celonis.de
Disco	Fluxicon	www.fluxicon.com

Thank you