COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

The Role of Process Mining for Financial Statement Audits

Michael Werner

Michael Werner

Associate Professor in Accounting Information Systems

Lecturer and Senior Lecturer for AIS

Researcher for AIS

Senior Consultant & Operations Manager

Research Focus:

- Data Science for Accounting and Auditing
- Data Analytics for Financial Statement Audits

Selected Appointments and Awards:

- Academic Consultant to the IAASB
- Chairman for the Outsourcing committee of the German Standardization Organization
- AIS Section Outstanding Dissertation Award from the American Accounting Association

Current Research Topics:

- Big Data and Accounting: A Critical Reflection on Audit Data Analytics and Technology Adoption in the Audit Profession
- Reshaping the Audit Process: Intelligent Data-Driven Audit Approaches
- The Impact of Process Mining on Financial Statement Audits: An Empirical Investigation

Doctor of Economics and Social Sciences (Dr. rer. pol.)

Master of Science in Information Systems

Master of Management

Certified Information Systems Auditor® (CISA)

Certified Information Security Manager® (CISM)

Certified in the Governance of Enterprise IT® (CGEIT)

Hobbies:

- Family, Sailing, Surfing, Beach Volleyball

Agenda

- Process Mining Basics
- Audit Standards and Novel Audit Data Analytics
- Integrating Process Mining into Financial Statement Audits
- Empirical Evidence and Real Life Examples
- Limitations for Using Process Mining in Financial Statement Audits
- Outlook to Deep Data Analytics, Automated Auditing and Artificial Intelligence Based on Process Mining

COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

Process Mining Basics

Motivation

[Deloitte Centre on Process Bionics](www.youtube.com/watch?v=6rnf3tw_6WY)

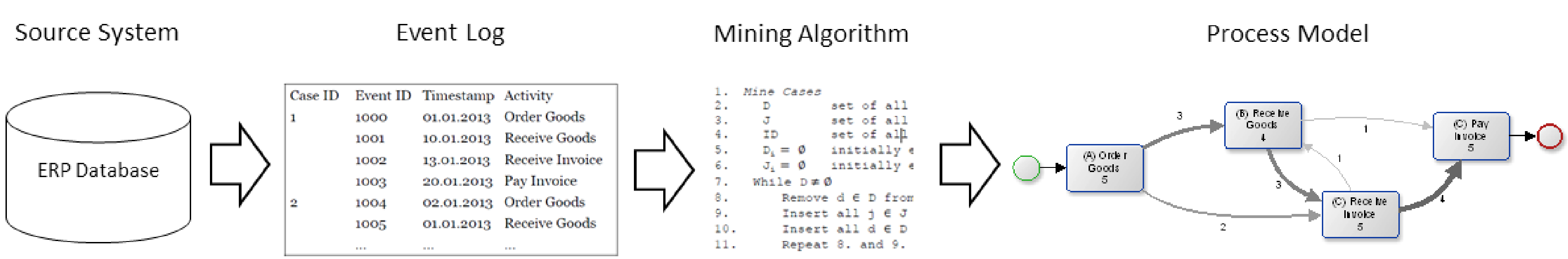
www.youtube.com/watch?v=6rnf3tw_6WY

Process Mining in a Nutshell

- The aim of process mining is the extraction of information about business processes (Van der Aalst 2011 p. 1).

- Process mining encompasses "techniques, tools and methods to discover, monitor and improve real processes (…) by extracting knowledge from event logs (…)" (Van der Aalst et al. 2012 p. 15).

- It is an unsupervised data mining technique for descriptive, predictive and prescriptive analytics.



Case ID	Event ID	Timestamp	Activity
1	1000	01.01.2013	Order Goods
	1001	10.01.2013	Receive Goods
	1002	13.01.2013	Receive Invoice
	1003	20.01.2013	Pay Invoice
2	1004	02.01.2013	Order Goods
	1005	01.01.2013	Receive Goods
	…	…	…

Process Mining Terminology

- **A process model** is an abstraction of the real world execution of a business process. A process model is a graphical representation of a business process that describes the dependencies between activities that need to be executed collectively for realizing a specific business objective. It consists of a set of activity models and constraints between them (Weske 2012 p. 7).

- Process models can be represented in different process **modeling languages** for example using the Business Process Model and Notation (BPMN), Event Driven Process Chains (EPC) or Petri Nets.

- An **event log** is a table that contains all recorded events that relate to executed business activities.

- Each **event** is mapped to a **case** in the event log.

- A single execution of a business process is called **process instance**. They are reflected in the event log as a set of events that are mapped to the same case. The model that describes the execution of a single process instance is called **process instance model**.

- The sequence of recorded events in a case is called **trace**.

- A process model abstracts from the single behavior of process instances and provides a model that reflects the behavior of all instances that belong to the same process.

- Cases and events are characterized by **classifiers** and **attributes**. Classifiers ensure the distinctness of cases and events by mapping unique names to each case and event. Attributes store additional information that can be used for analysis purposes.

Process Mining Algorithms

Deterministic mining algorithms

- Deterministic mining algorithms produce defined and reproducible results.
- *Examples: the α-Algorithm (van der Aalst and van Dongen 2002), MLPM algorithm (Werner and Gehrke 2015)*

Heuristic mining algorithms

- Deterministic algorithms that incorporate frequencies of events and traces for reconstructing a process model
- *Examples: Heuristic Miner (Weijters et al. 2006)*

Genetic mining algorithms

- Non-deterministic algorithms that mimic the process of natural evolution: initialization, selection, reproduction and termination
- *Example: Genetic Process Miner (de Medeiros 2006)*

Event Logs

Classifier

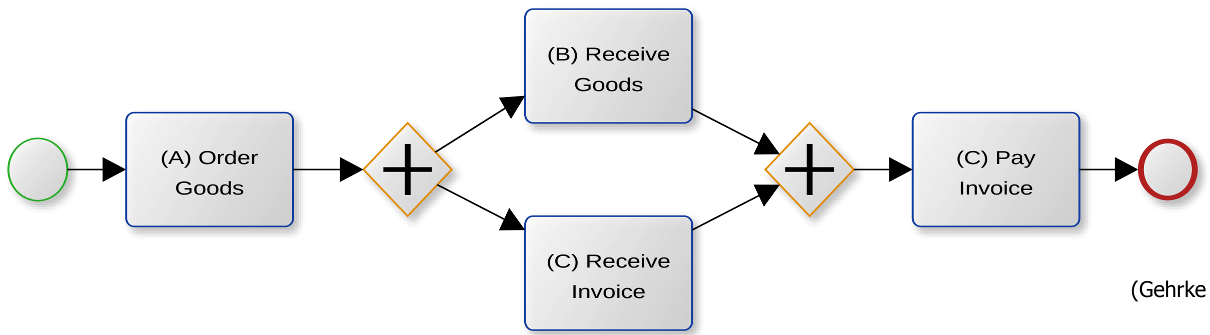
Attributes

Case ID	Event ID	Timestamp	Activity	Resource
1	1000	01.01.2013	(A) Order Goods	Peter
	1001	10.01.2013	(B) Receive Goods	Michael
	1002	13.01.2013	(C) Receive Invoice	Frank
	1003	20.01.2013	(D) Pay Invoice	Tanja
2	1004	02.01.2013	(A) Order Goods	Peter
	1005	03.02.2013	(B) Receive Goods	Michael
	1006	05.02.2013	(C) Receive Invoice	Frank
	1007	06.02.2013	(D) Pay Invoice	Tanja
3	1008	01.01.2013	(A) Order Goods	Louise
	1009	04.01.2013	(C) Receive Invoice	Frank
	1010	05.01.2013	(B) Receive Goods	Michael
	1011	10.01.2013	(D) Pay Invoice	Tanja
4	1016	15.01.2013	(A) Order Goods	Peter
	1017	20.01.2013	(C) Receive Invoice	Claire
	1018	25.01.2013	(D) Pay Invoice	Frank
5	1023	01.01.2013	(A) Order Goods	Michael
	1024	10.01.2013	(B) Receive Goods	Michael
	1025	13.01.2013	(C) Receive Invoice	Michael
	1026	20.01.2013	(D) Pay Invoice	Michael

(Gehrke & Werner al. 2013, p. 937)

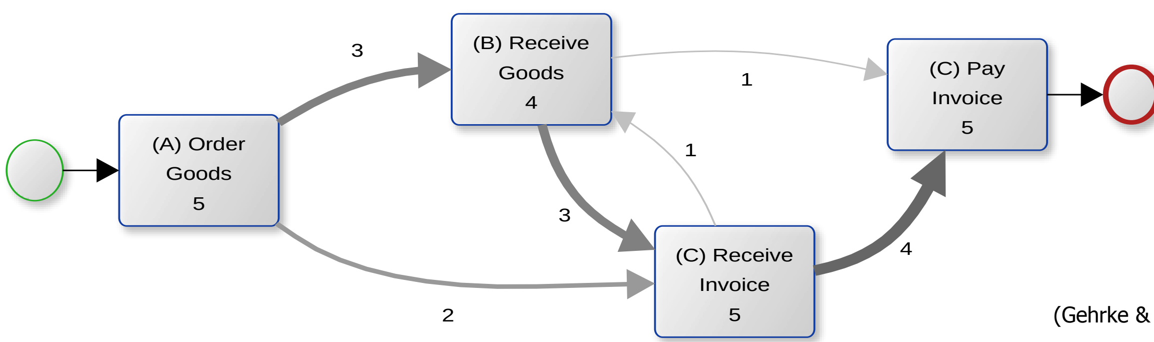
Mining Output

- The α-Algorithm produces a process model with concurrent activities B and C



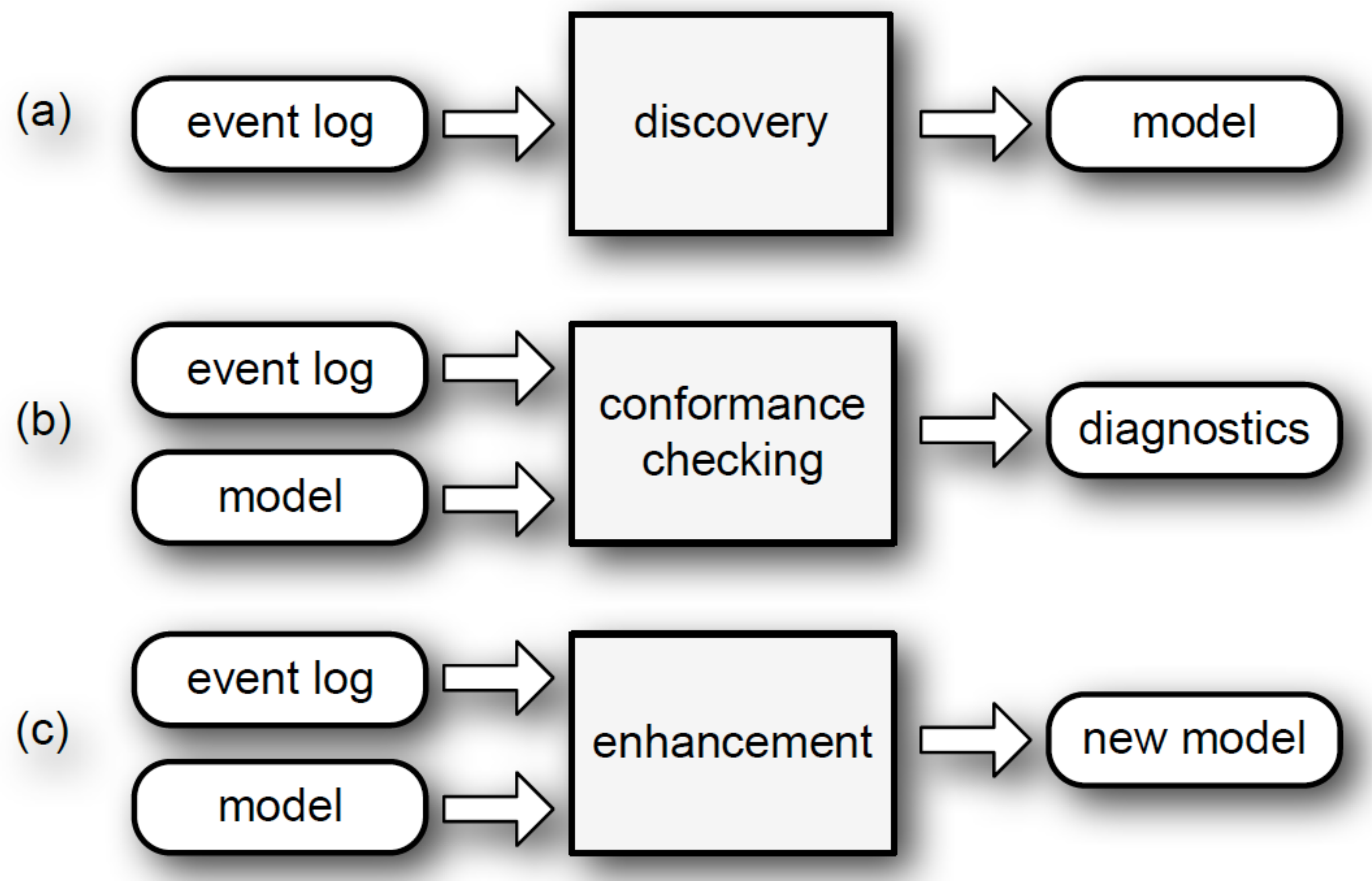
(Gehrke & Werner al. 2013, p. 938)

- The Fuzzy Miner algorithm produces a dependency graph, the thickest paths illustrate the "highways" in the process model



(Gehrke & Werner al. 2013, p. 939)

Application Areas



(van der Aalst et al. 2012, p. 4)

Audit Standards and Novel Audit Data Analytics

Core Challenge for Financial Statement Audits

Company Perspective

System based und automated transaction processing

- The financial reporting is prepared based on data that is created in ERP systems during transaction processing.
- Transactions are processed automated by ERP systems creating increasingly large data amounts.

Discrepancy

Auditor Perspective

Manual audit procedures

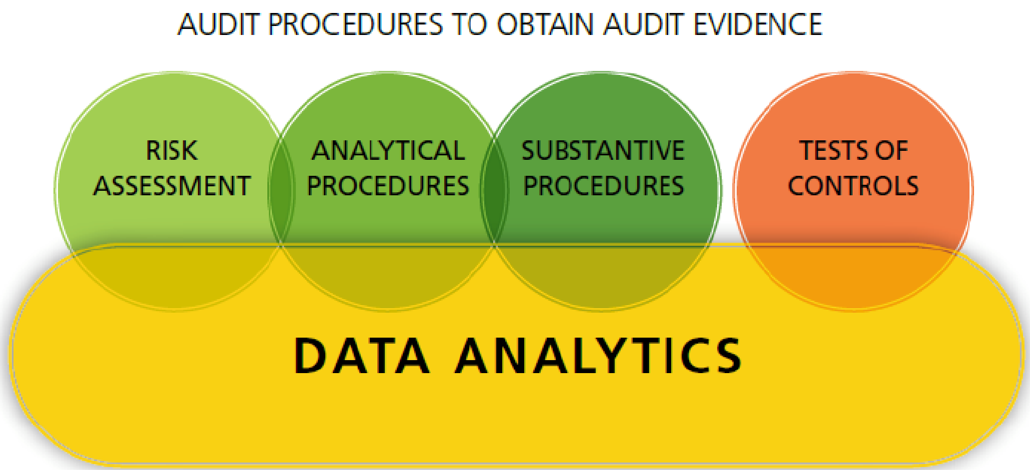
- Has to issue an opinion about the truth and fairness of the financial statements.
- Audits primarily based on traditional and manual audit procedures.

Problem:

- Imbalance between automated processing on the companies' side and manual audit procedures on the auditors' side
- ➢ Inefficient or even ineffective audits

Audit Standards and Data Analytics

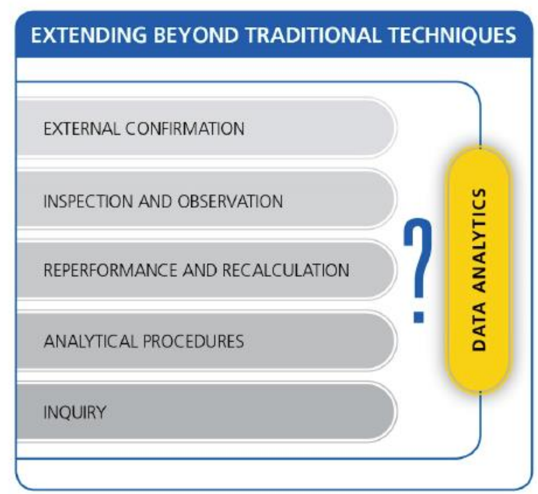
- Insecurity in the profession
  - Can novel data analysis techniques be used?
  - How can they be embedded into contemporary FSA?
- Audit standards neither encourage the use of innovative audit data analytics nor prohibit it.



(IFAC 2016, p. 7)

- High regulatory pressure in the profession (oversight bodies, peer-reviews, personal liability etc.).

➢ Little innovation, sticking to traditional audit procedures which will pass peer-reviews easily (checkmark mentality).



(IFAC 2016, p. 10)

IAASB Data Analytics Working Group

- Set up to *exploring the growing use of technology* in the audit, with a *focus on data analytics*
- Request for input released in September 2016
- Selected **Key Aspects**:
  - Increasingly complex environments
  - Higher transaction volumes, increasingly complex operations
  - Concept of CAATs (Computer Assisted Audit Techniques) was created in a completely different technological era
  - Technological change at a rapid pace
  - Changing stakeholder expectations regarding the use of technology
  - FSA can be enhanced by the use of data analytics
  - ADA will not replace *professional judgement* and *professional skepticism*
- Selected **Open Questions**:
  - How does an engagement team classify the audit evidence provided by data analytics?
  - What is an appropriate level of work effort for identified exceptions (now 100% of population is tested)?
  - What are the implications for small and medium practices?
  - How to deal with legacy systems especially for public sector entities?

➢ Input to revised version of ISA 315 “Identifying and Assessing the Risks of Material Misstatement through Understanding the Entity and Its Environment”

  ➢ Explicit recognition of importance and the use of automated tools and techniques to perform risk assessment procedures

➢ Revision of ISA 520 “Analytical Procedures” or release of a non-authoritative guidance

Integrating Process Mining into Financial Statement Audits

COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

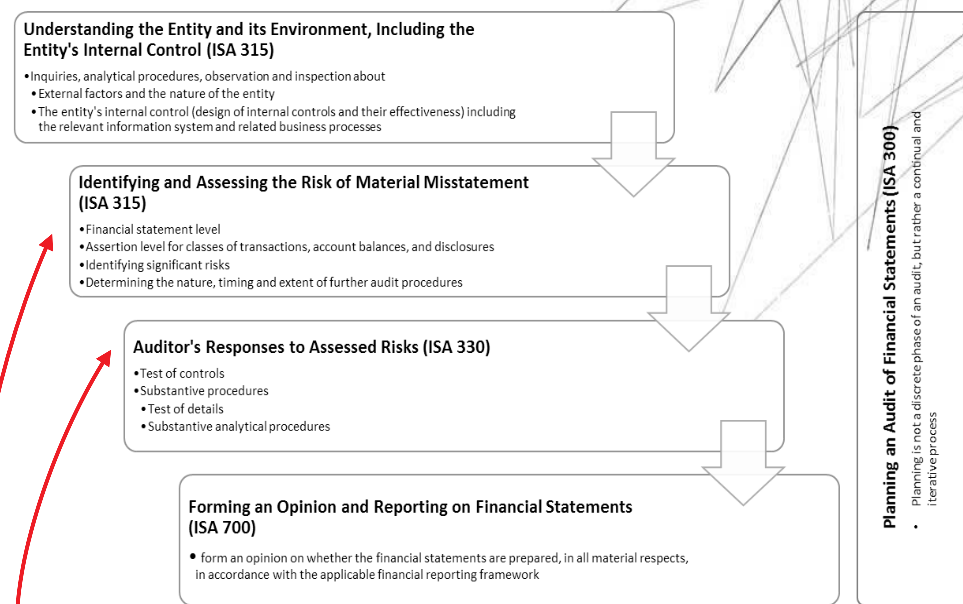
Business Processes and Financial Statement Audits

- Well-controlled business processes lead to correct entries in the financial accounts.
- The audit of internal controls systems is much more efficient than substantial audit procedures.

Traditional audit procedures like
- Interviews
- Inspection of single transactions

are not efficient for
- High transaction volumes
- Complex business processes
- Decreasing transparency due to the growing automation by the integration of information systems

Financial Statement Audit Process



(Werner 2019, p. 1048)

Pre-condition for these phases:

- **Reliable understanding of relevant business processes** especially related to:
  - Relationship between business processes and financial accounts
  - Internal controls

CBS COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

Integrating Process Mining into Contemporary Audits

Audit Phase	Supported Audit Activities	Relevant Audit Standard
Understand the entity	Obtain an understanding of the composition of the balance sheet and income statement accounts to establish an audit strategy	ISA 240.22, ISA 300.2, ISA 315.6
	Initial assessment if reliance on the internal controls system is feasible.	ISA 315.20, ISA 315.26, ISA 330.6
	Documentation of obtained understanding regarding the audited entity.	ISA 315.32, ISA 230.8, ISA 330.28
Identify & assess risks	Obtain an understanding of the significant classes of transactions related to the analyzed processes.	ISA 315.18, ISA 315.20, ISA 315.26
	Determination if the processing of significant classes of transactions is highly automated.	ISA 315.30, ISA 330.8
	Assessment of relevant risks via automated analysis of identified business processes for each relevant assertion and each significant account or disclosure.	ISA 315.5, ISA 315.6, ISA 315.25
	Assessment if control activities are implemented in the analyzed processes for significant risks and whether these have been designed and implemented to achieve relevant control objectives.	ISA 315.13, ISA 315.20, ISA 315.21 ISA 315.29
Design and execute responses to risks	Identification of activities in the analyzed process that represent control activities relevant to the audit. A control exception is indicated if such an activity was missing in a process variant to a material extent.	ISA 330.6, ISA 330.15
	Assessment when and by whom a control activity was executed. Evaluating the operating effectiveness of controls relevant to the audit by directly analyzing the recorded transaction data values within a process.	ISA 330.8, ISA 330.10
	Identification of deviations from standard procedures to assess potential impact on the reliance of internal controls and to guide further substantive audit procedures.	ISA 330.17
Conclude and communicate	Documentation of the number of transactions and their total posted amount that is affected by a control exception.	ISA 230.2, ISA 315.20, ISA 330.17
	Assessment whether the control exception rate exceeds acceptable tolerances.	ISA 530.13
	Modification of risk assessment if the types of exceptions is systematic and no compensating controls can be identified.	ISA 315.20, ISA 330.17
	Observation and communication of possible root causes of control exceptions to the auditee.	ISA 265.5, ISA 330.17, ISA 530.12

Testing Internal Controls via Process Mining

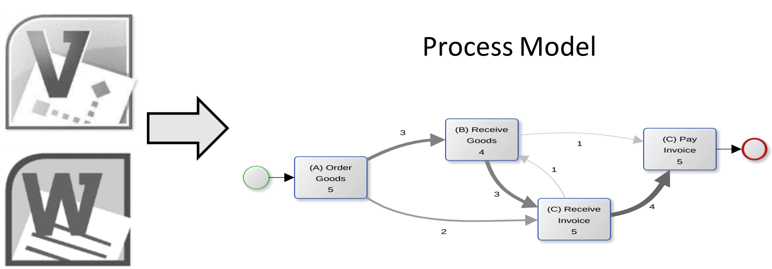
Traditional Test of Controls

Automated Test of Controls via Process Mining

Data Collection via Interviews

Manual Process Modelling

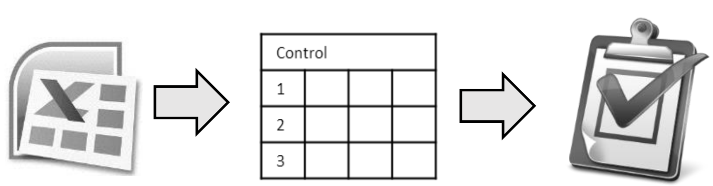
Process Model



Manual Testing of Internal Controls

Control Matrix

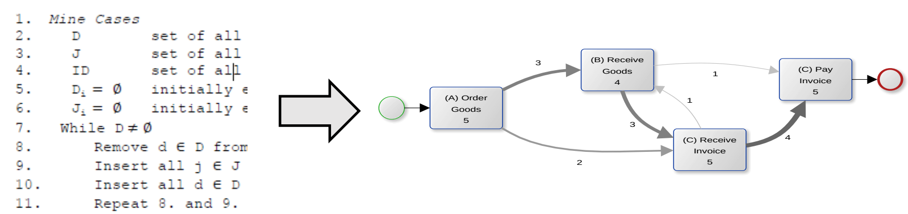
Report



Automated Data Extraction

Case ID	Event ID	Timestamp	Activity
1	1000	01.01.2013	Order Goods
	1001	10.01.2013	Receive Goods
	1002	13.01.2013	Receive Invoice
	1003	20.01.2013	Pay Invoice
2	1004	02.01.2013	Order Goods
	1005	01.01.2013	Receive Goods
	...	...	...

Automated Process Discovery



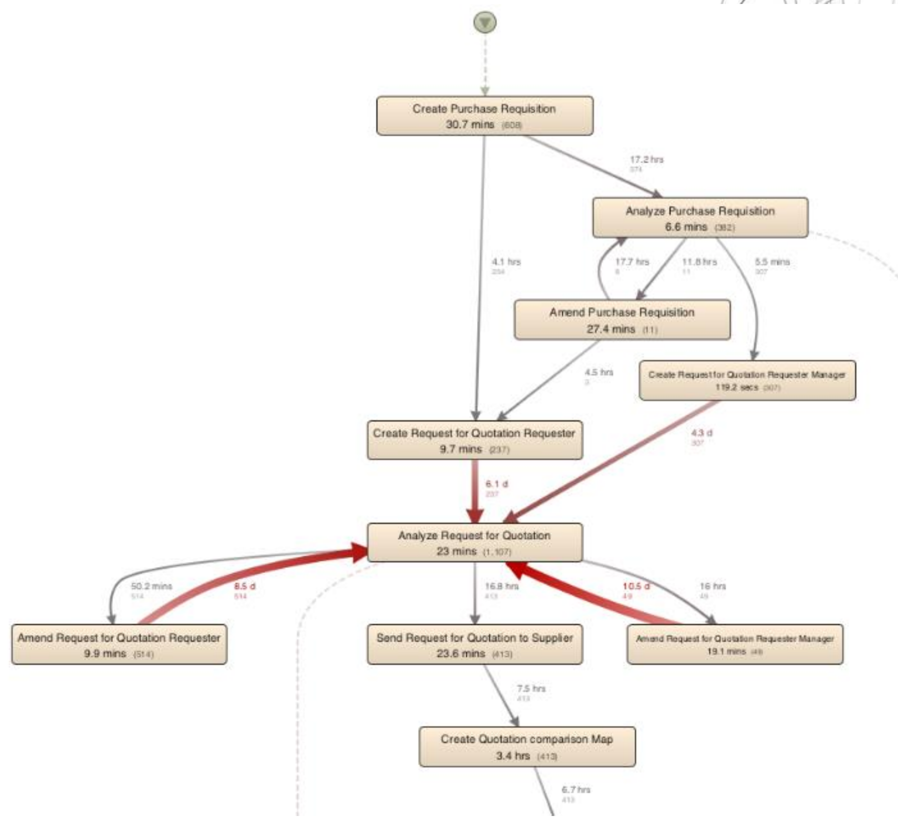
Automated Test of Controls

(adapted from Werner & Gehrke 2018, p. 37)

Empirical Evidence and Real Life Examples

Example of a Mined Process Model for a Procurement Process

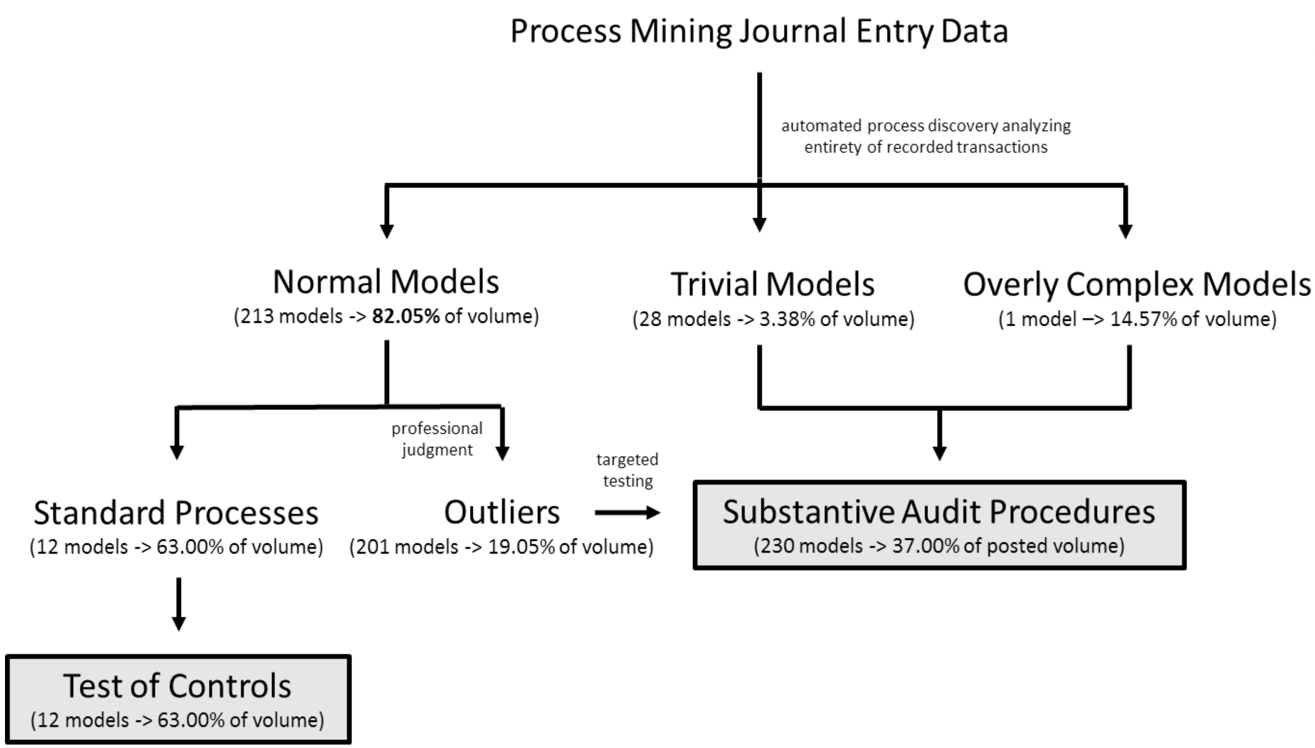
- What is the average time for:
    - *Manually modelling* a business process of moderate complexity in an internationally operating company during a financial statement audit?



Created using Disco (fluxicon 2018)

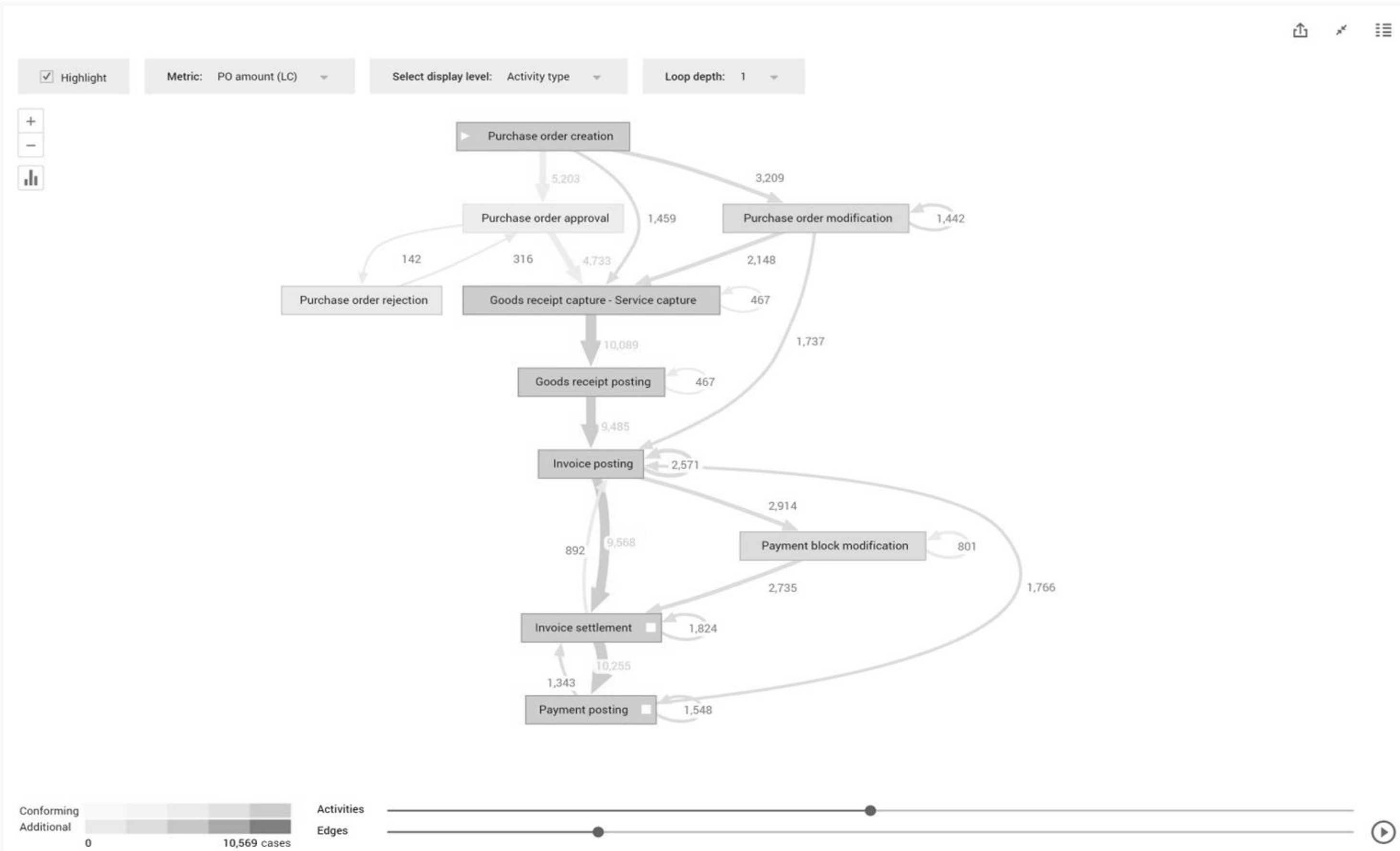
Empirical Evidence

- Study based on 3 real SAP data sets.
- Up to 82% of the overall posted volume in the financial accounts can be analyzed automatically via process mining techniques.



Example: Process Mining for Testing Internal Controls of a Procurement Process

Mined Process Model



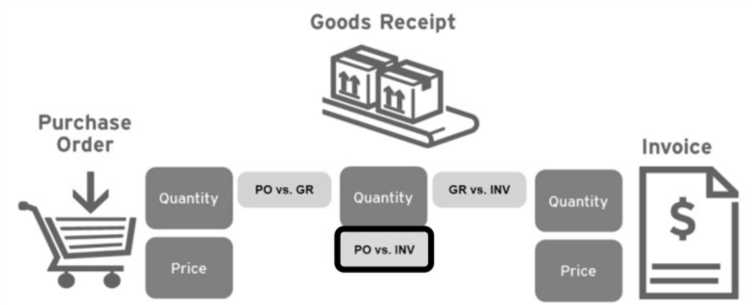
Example: Process Mining for Testing Internal Controls of a Procurement Process

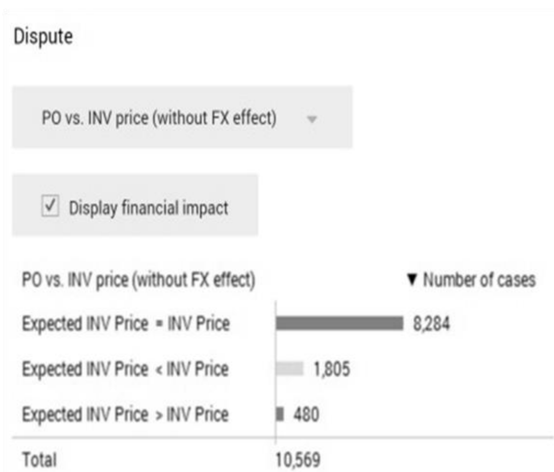
Financial Reconciliation - Account Coverage - Account activity coverage through process mining

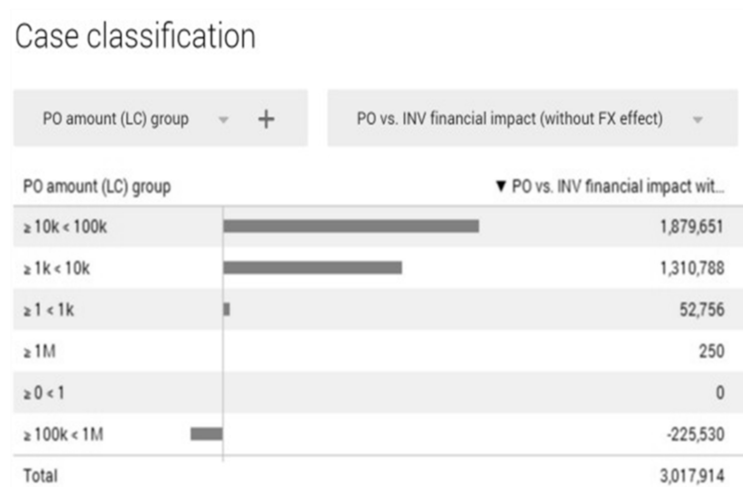
GL account description	Coverage	▼ Debit	Credit	Net activity (LC)
GR/IR clearing account for trading goods	Covered	74,709,931 (90.63%)	73,032,804 (90.43%)	1,677,127
	Not covered	7,727,891 (9.37%)	7,727,891 (9.57%)	0
Trade accounts payable-third Local	Covered	70,715,899 (95.27%)	71,858,524 (95.14%)	-1,142,625
	Not covered	3,513,415 (4.73%)	3,667,892 (4.86%)	-154,478
GR/IR clearing-others	Not covered	44,384,744 (99.99%)	46,061,872 (99.99%)	-1,677,127
	Covered	2,364 (0.01%)	2,332 (0.01%)	33
Trade accounts payable	Not covered	17,754,766 (81.25%)	15,783,260 (67.90%)	1,971,506
	Covered	4,098,103 (18.75%)	7,462,149 (32.10%)	-3,364,045
Trade a/cs - Third Manual	Not covered	15,006,336 (100.00%)	15,710,487 (100.00%)	-704,151

Example: Process Mining for Testing Internal Controls of a Procurement Process

Automated Control Testing (Two-Way-Match)



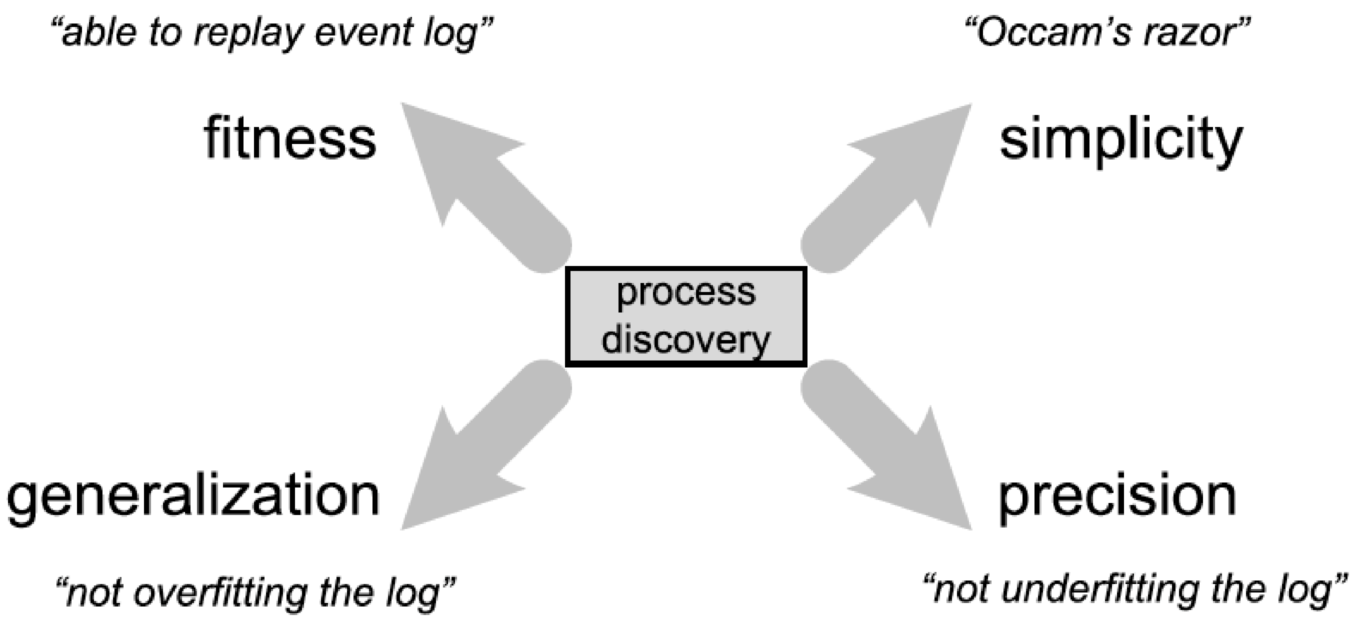




Limitations for Using Process Mining in Financial Statement Audits

COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

Competing Quality Criteria



(van der Aalst 2016, p. 189)

Fitness

- Ability of a model to replay all behavior recorded in the event log.

Simplicity

- The simplest model that can explain the observed behavior should be preferred.

Precision

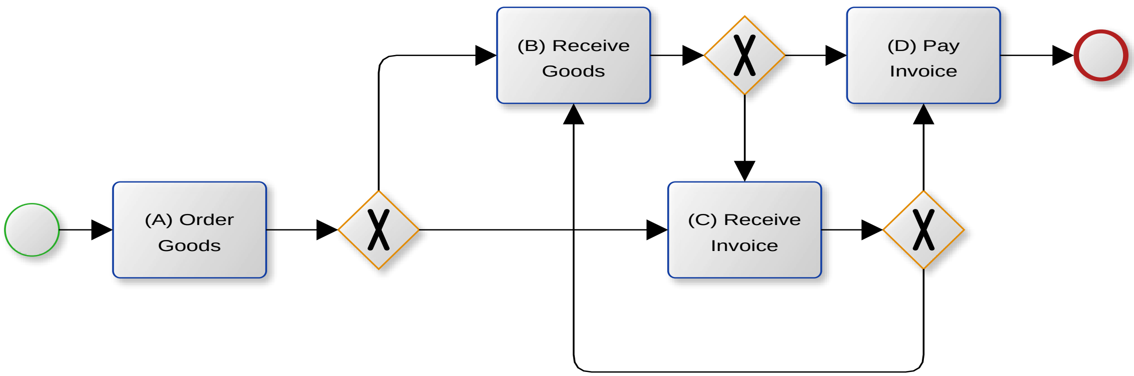
- The model does not allow additional behavior very different from the behavior recorded in the event log.

Generalization

- A process model is not exclusively restricted to display the eventually limited record of observed behavior in the event log.

Audit Implications

- The mined model has a high fitness but poor precision.
- *All observed traces can be replayed*, but it also allows for many (infinitive) *more traces*.



(Gehrke & Werner al. 2013, p. 938)

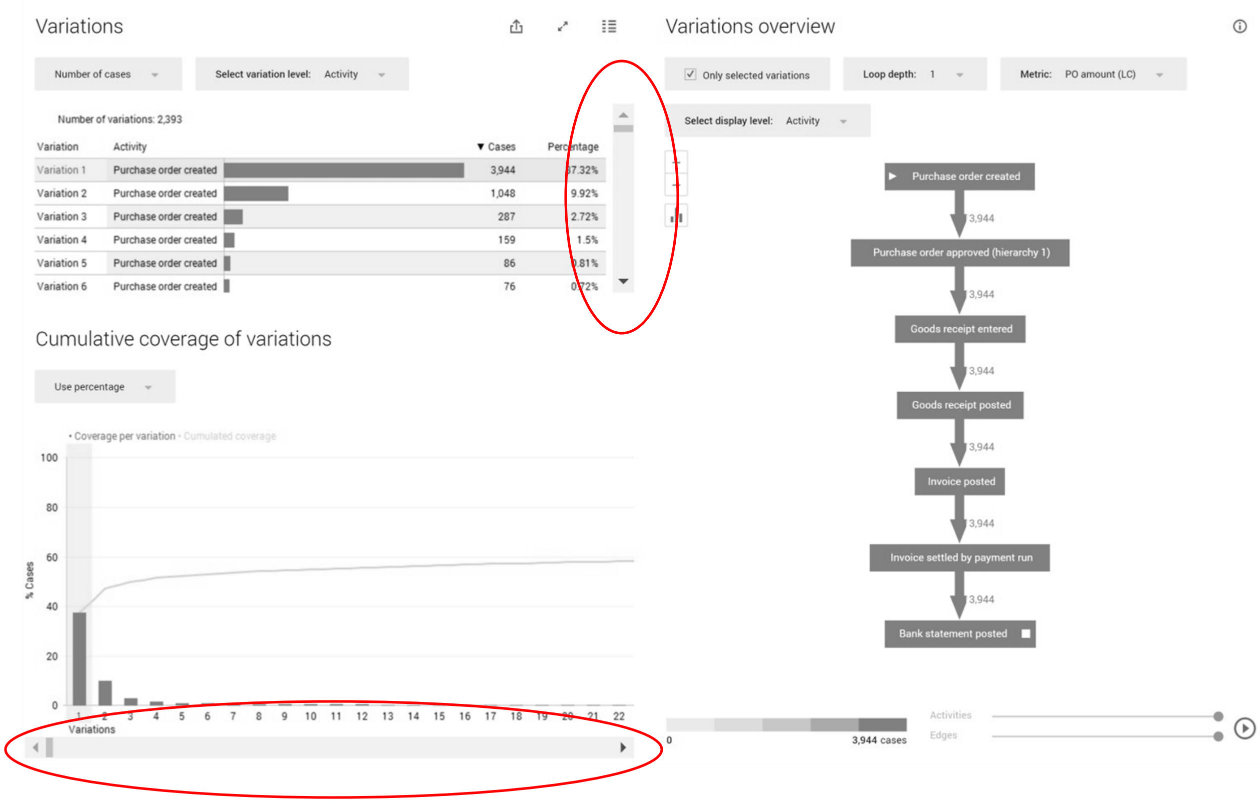
- ➢ Unfitting model -> *False negative* audit results (compliance violations are not detected)
- ➢ Imprecise model -> *False positive* audit results (compliance violations are indicated that did not occur in reality)

Sample Event Log

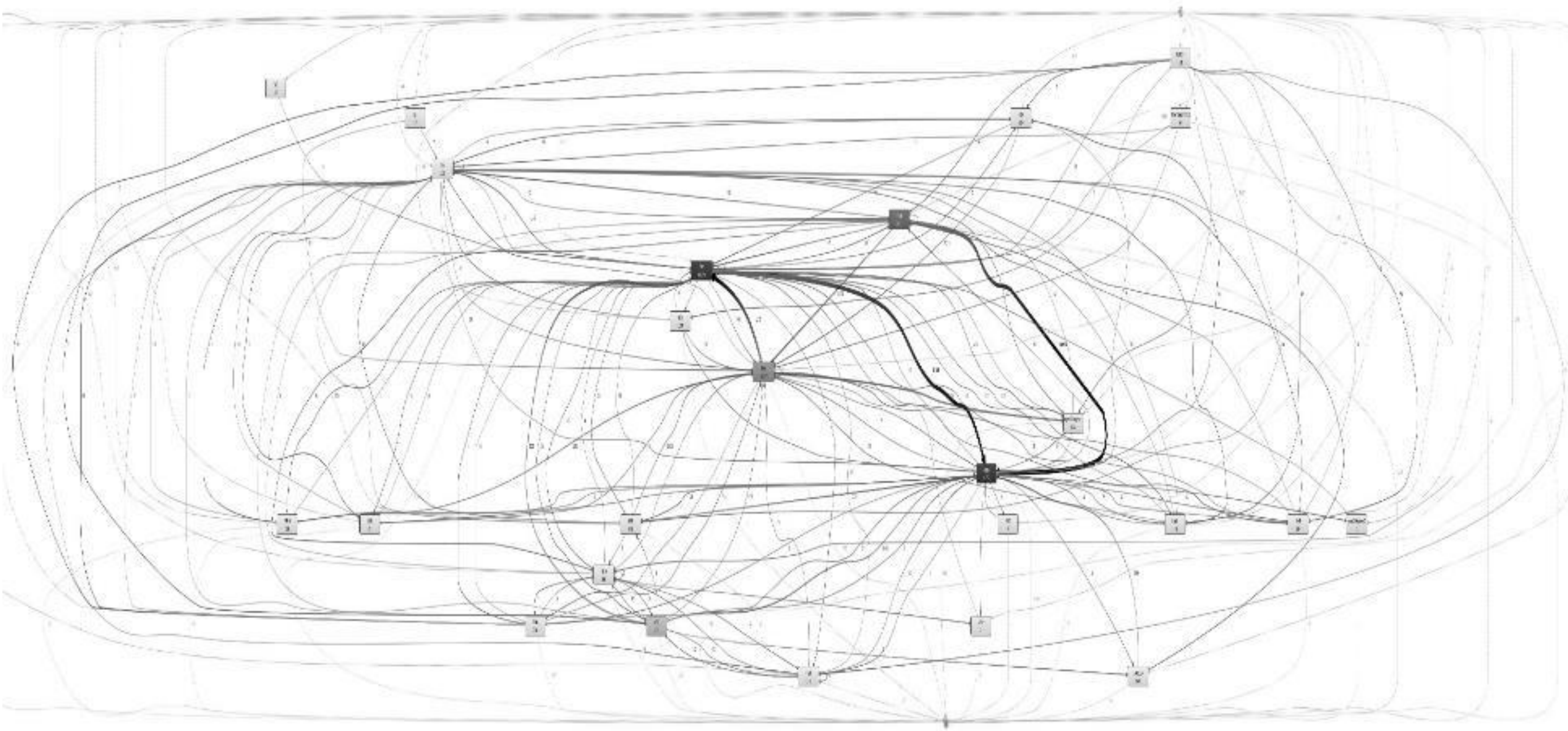
Case ID	Event ID	Timestamp	Activity
1	1000	01.01.2013	(A) Order Goods
	1001	10.01.2013	(B) Receive Goods
	1002	13.01.2013	(C) Receive Invoice
	1003	20.01.2013	(D) Pay Invoice
2	1004	02.01.2013	(A) Order Goods
	1005	03.02.2013	(B) Receive Goods
	1006	05.02.2013	(C) Receive Invoice
	1007	06.02.2013	(D) Pay Invoice
3	1008	01.01.2013	(A) Order Goods
	1009	04.01.2013	(C) Receive Invoice
	1010	05.01.2013	(B) Receive Goods
	1011	10.01.2013	(D) Pay Invoice
4	1016	15.01.2013	(A) Order Goods
	1017	20.01.2013	(C) Receive Invoice
	1018	25.01.2013	(D) Pay Invoice
5	1023	01.01.2013	(A) Order Goods
	1024	10.01.2013	(B) Receive Goods
	1025	13.01.2013	(C) Receive Invoice
	1026	20.01.2013	(D) Pay Invoice

Process Variants

High number of process variations and deviations



Process Complexity and Spaghetti Models



Maturity of Event Log Data

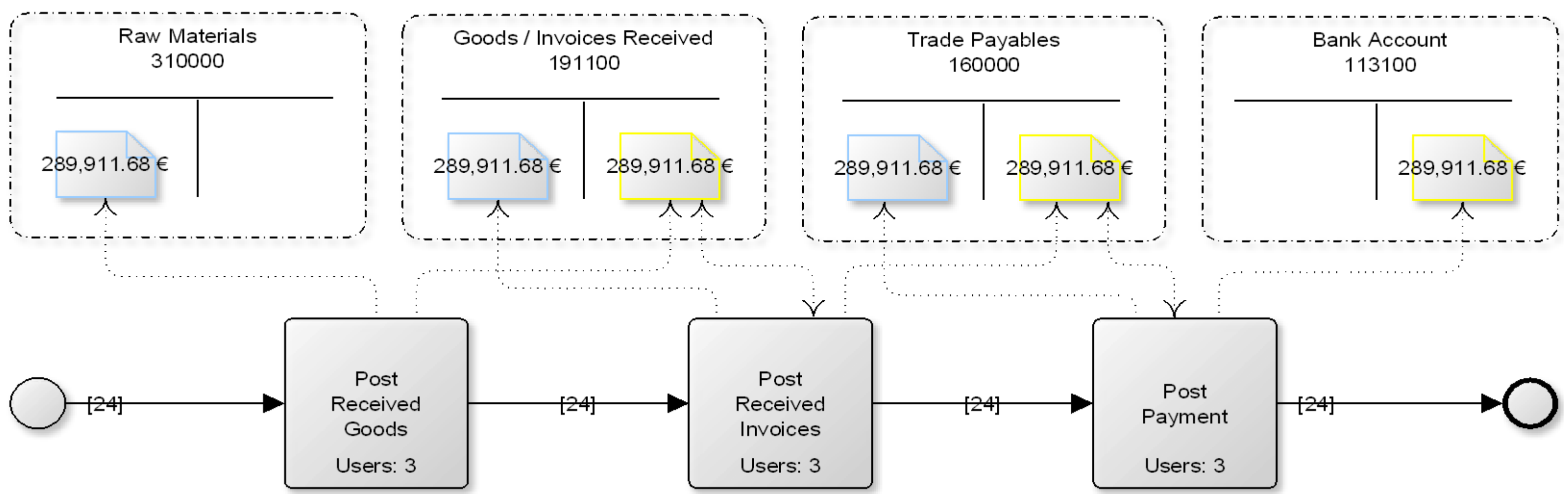
The quality of an event log depends on the source system's ability to record process relevant data.

Level	Characterization	Examples
★★★★★	Highest level: the event log is of excellent quality (i.e., trustworthy and complete) and events are well-defined. Events are recorded in an automatic, systematic, reliable, and safe manner. Privacy and security considerations are addressed adequately. Moreover, the events recorded (and all of their attributes) have clear semantics. This implies the existence of one or more ontologies. Events and their attributes point to this ontology.	Semantically annotated logs of BPM systems.
★★★★	Events are recorded automatically and in a systematic and reliable manner, i.e., logs are trustworthy and complete. Unlike the systems operating at level ★★★, notions such as process instance (case) and activity are supported in an explicit manner.	Events logs of traditional BPM/ workflow systems.
★★★	Events are recorded automatically, but no systematic approach is followed to record events. However, unlike logs at level ★★, there is some level of guarantee that the events recorded match reality (i.e., the event log is trustworthy but not necessarily complete). Consider, for example, the events recorded by an ERP system. Although events need to be extracted from a variety of tables, the information can be assumed to be correct (e.g., it is safe to assume that a payment recorded by the ERP actually exists and vice versa).	Tables in ERP systems, event logs of CRM systems, transaction logs of messaging systems, event logs of high-tech systems, etc.
★★	Events are recorded automatically, i.e., as a by-product of some information system. Coverage varies, i.e., no systematic approach is followed to decide which events are recorded. Moreover, it is possible to bypass the information system. Hence, events may be missing or not recorded properly.	Event logs of document and product management systems, error logs of embedded systems, worksheets of service engineers, etc.
★	Lowest level: event logs are of poor quality. Recorded events may not correspond to reality and events may be missing. Event logs for which events are recorded by hand typically have such characteristics.	Trails left in paper documents routed through the organization ("yellow notes"), paper-based medical records, etc.

(van der Aalst et al. 2012, p. 7)

Integration of Financial Information

- A challenge in using process mining in financial statement audits is the integrations of key information concepts:
  - Financial accounts
  - Internal controls
  - Materiality
  - Etc.



(Werner 2019, p. 1049)

Further Challenges

System Boundaries

- Only data recorded in the source systems can be analyzed.

Noise and Incompleteness

- Noise refers to rare and infrequent behavior.
- Incompleteness means that not all possible behavior is recorded.

Data Extraction

- Extracting, transforming and loading the source data is a non-trivial process seldom supported by process mining tools.

Concept Drift

- Business processes change over time which is not considered by contemporary process mining tools.

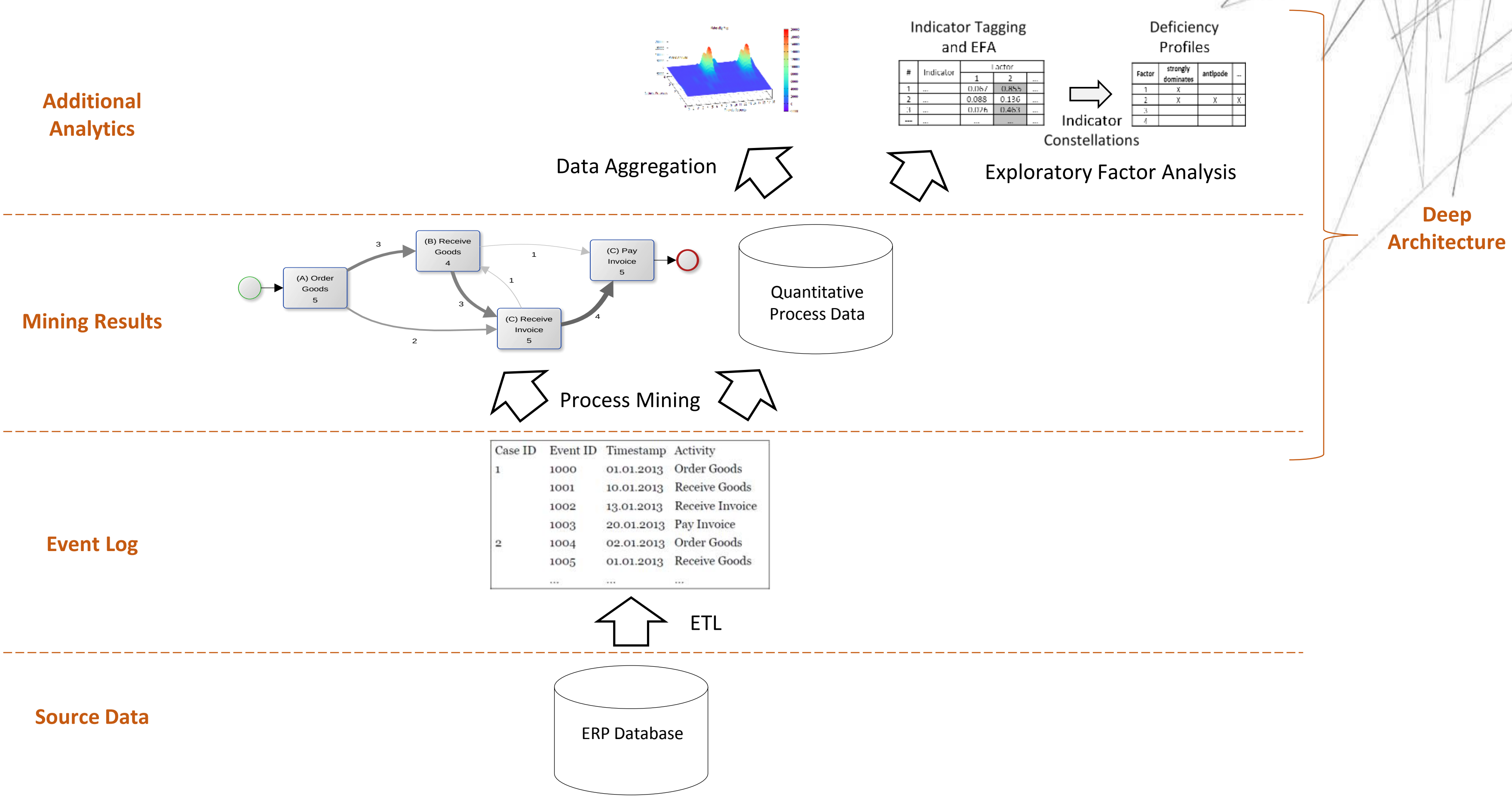
Data Privacy

- Process mining can easily be used to measure employee performance on an activity basis which leads to ethical questions and legal implications.

Outlook to Deep Data Analytics, Automated Auditing and Artificial Intelligence Based on Process Mining

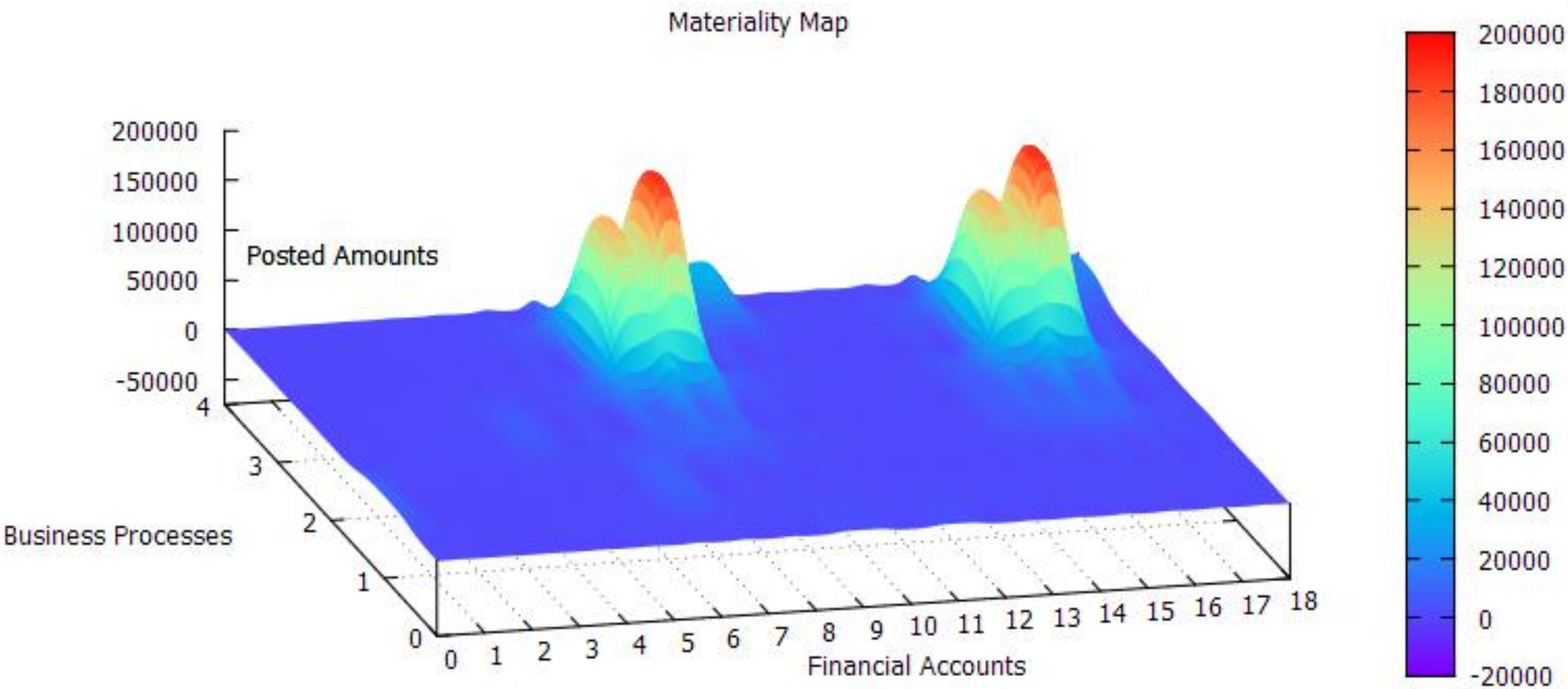
CBS COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

Deep Data Analytics



Aggregating Mining Data into Materiality Process Maps

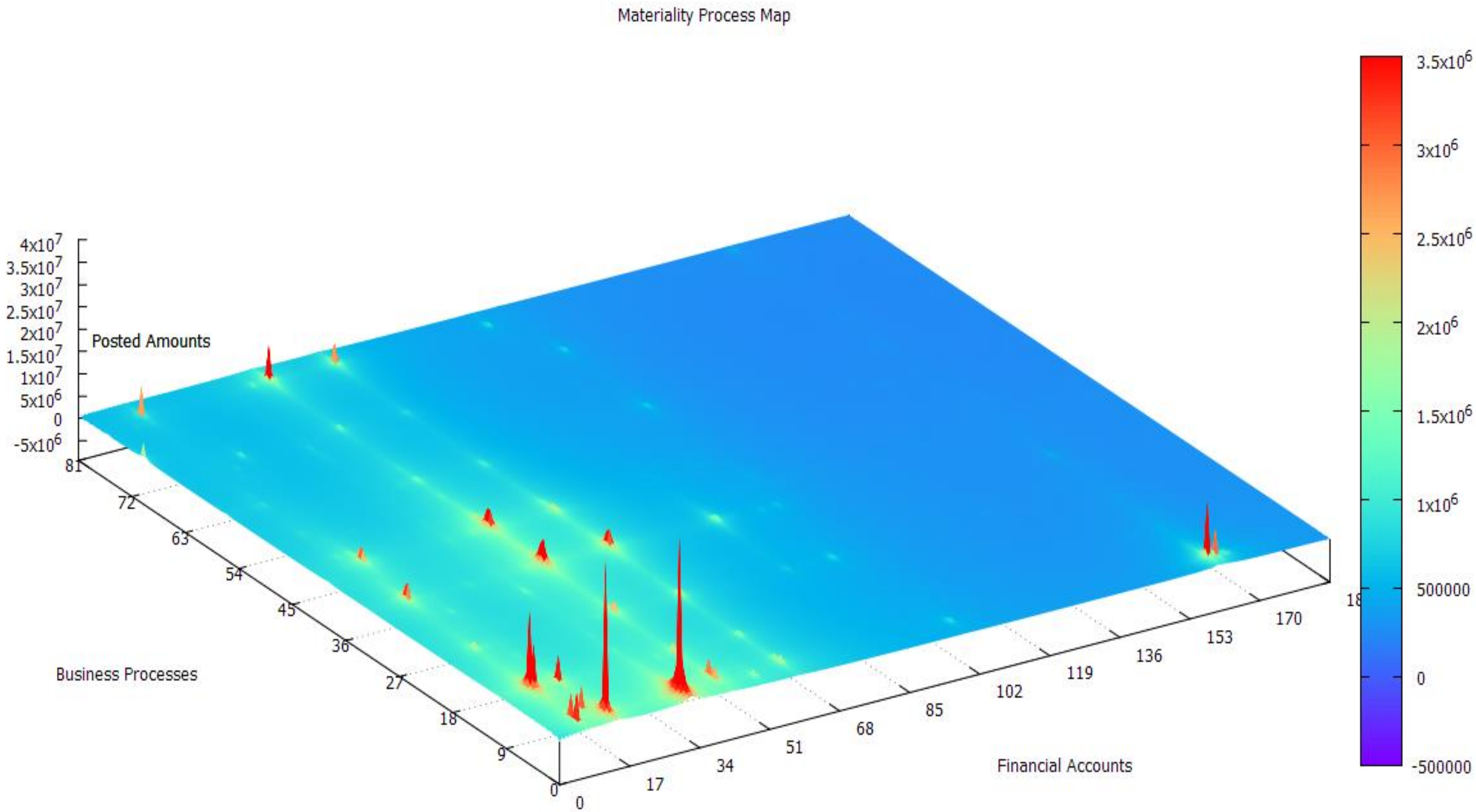
Example refers to a set of five simple process models mined from the SAP IDES test data.



Die diagram shows the amount of journal postings (ordinate) per mined process (applicate) and effected financial accounts (abscissa).

(Werner 2019, p. 1050)

Materiality Process Map for IDES Test Data Set



(Werner 2019, p. 1050)

- 82 processes that met these specifications. They caused postings in 188 different accounts.
- Trivial processes and those with total postings less then 1 million € were neglected.

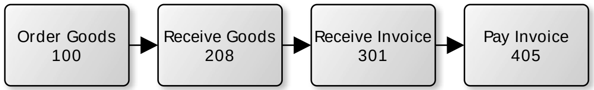
Detecting Internal Control Deficiency Profiles via Exploratory Factor Analysis

1. Mining process models for each process execution (case).

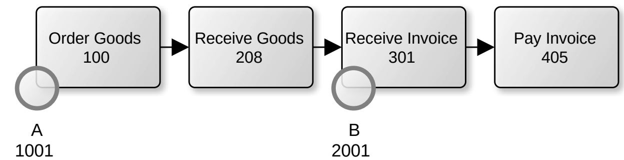
2. Analyze each model for critical data constellations (indicators).

3. Statistically analyze tagged models to find patterns that indicate flaws in the internal control system.

25,051 mined models



9,218 tagged models



#	Variables	Factors		
		1	2	…
1	A transaction is executed with administrator access rights	-0.067	**-0.855**	…
	…			
3	An accounts payable subsidiary ledger account is adjusted using the transaction to record an incoming invoice and then concealed via general ledger entries by the same user	-0.026	**0.463**	-0.106
4	An incoming invoice is created and payment for it initiated by the same user	-0.020	**0.839**	…
	…			
13	The customer master data is changed and an unauthorized invoice entered by the same user	**0.713**	-0.033	…
	…			
16	An accounts receivable subsidiary ledger account is adjusted using payment runs and this is then concealed with general ledger entries	**0.688**	0.041	…
17	A sales invoice is changed and payments are changed for it by the same user	**0.876**	-0.015	-..
	…			

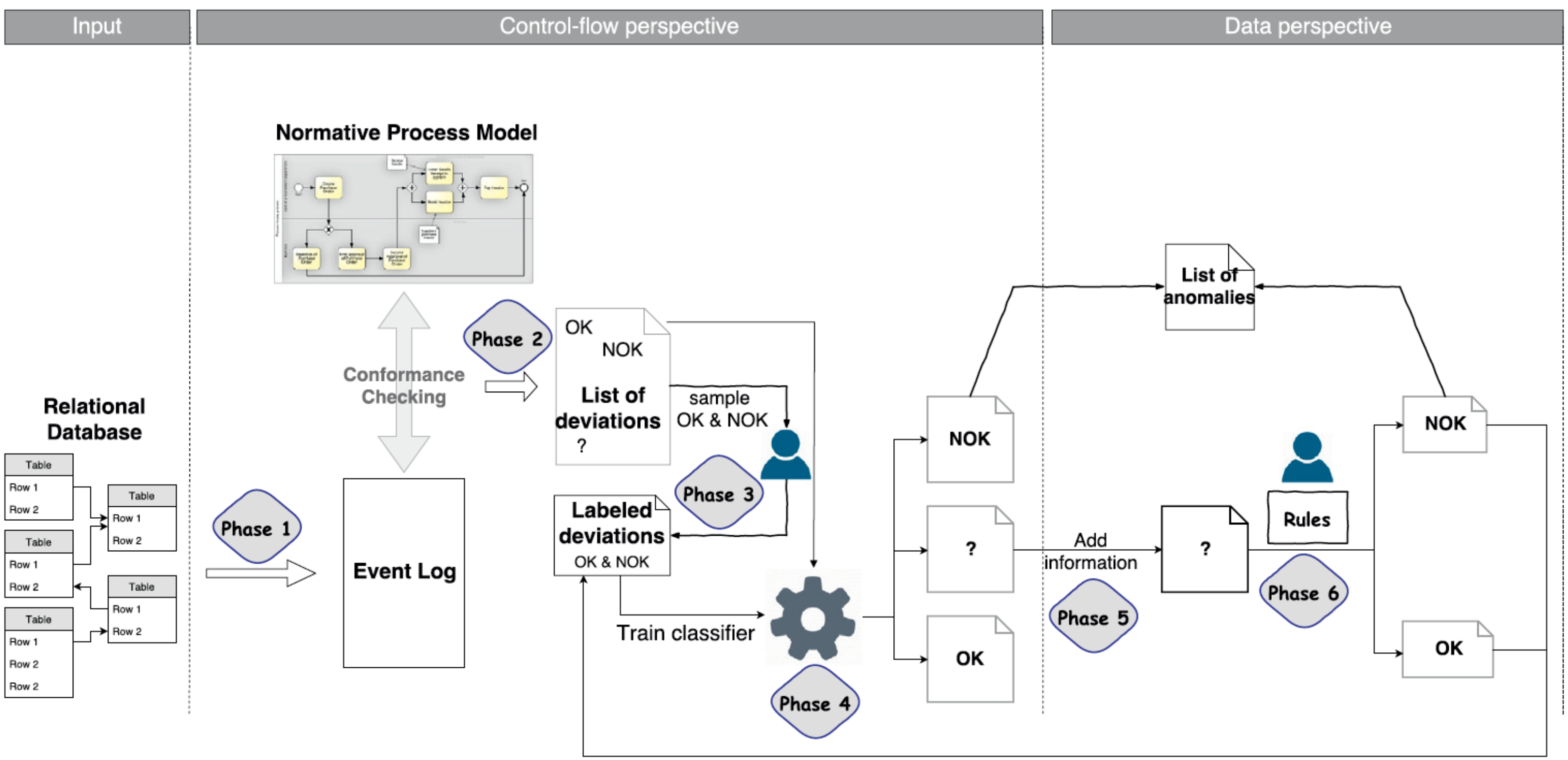
Systemic internal control system deficiency in the procurement process

Systemic internal control system deficiency in the sales process

(Werner & Gehrke 2018, p. 43-47)

Integration of Artificial Intelligence for Classifying Deviations

- Mining results can indicate a flood of deviations.
- Supervised classification algorithms can be used to reduce the number of false positives.



(Jans & Hosseinpour 2018, p. 9)

References

- Aalst, Wil M. P. van der. *Process Mining: Data Science in Action*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016.
- Aalst, Wil M. P. van der, and B.F. van Dongen. "Discovering Workflow Performance Models from Timed Logs." In *Engineering and Deployment of Cooperative Information Systems: First International Conference, Proceedings*, edited by Yanbo Han, Stefan Tai, and Dietmar Wikarski, 45–63. Beijing: Springer, 2002.
- Aalst, Wil M. P. van der, A. Andriansyah, Alves K. de Medeiros, F. Arcieri, T. Baier, T. Blickle, J. C. Bose, P. van den Brand, R. Brandtjen, and J. Buijs. "Process Mining Manifesto." In *BPM 2011 Workshops Proceedings*, 169–194, 2012.
- fluxicon. "Process Mining and Process Analysis - Fluxicon," 2018. www.fluxicon.com.
- Gehrke, Nick, and Michael Werner. "Process Mining." *Wisu - Das Wirtschaftsstudium*, no. 7 (2013): 934–43.
- Medeiros, A. K. Alves de. "Genetic Process Mining." Eindhoven University of Technology, 2006.
- Weijters, A., Wil M. P. van der Aalst, and A. K. A. de Medeiros. "Process Mining with the Heuristics Miner-Algorithm." *Technische Universiteit Eindhoven, Tech. Rep. WP* 166 (2006).
- Werner, Michael, and N. Gehrke. "Multilevel Process Mining for Financial Audits." *IEEE Transactions on Services Computing* 8, no. 6 (2015): 820–32.
- Werner, Michael, and N.Gehrke. "Indicator-Based Audit Procedures – An Alternative Approach for Internal Control Audits." *Journal of Information Systems*, (2018) Early Online.
- Werner, Michael. "Materiality Maps – Process Mining Data Visualization for Financial Audits." In *Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS 2019)*, 1045–54. Maui: IEEE, 2019.
- IFAC. "Exploring the Growing Use of Technology in the Audit, with a Focus on Data Analytics," September 2016.
- Jans, Mieke, and M.Hosseinpour. "How Active Learning and Process Mining Can Act as Continuous Auditing Catalyst." *International Journal of Accounting Information Systems*, December 2018.

Thank you for your attention

I am working closely together with process mining software companies and audit firms that are experts in process mining.

Please do not hesitate to contact me, if you have any questions.

DR MICHAEL WERNER

Associate Professor in Accounting Information Systems

CISM, CISA, CGEIT

COPENHAGEN BUSINESS SCHOOL

DEPARTMENT OF ACCOUNTING

Room C4.08

Solbjerg Plads 3

2000 Frederiksberg

Denmark

Tel.: +45 3815 2335

mwe.acc@cbs.dk

CBS COPENHAGEN BUSINESS SCHOOL HANDELSHØJSKOLEN