Annual Meeting of the IEEE Task Force on Process Mining

Monday September 19th, 2016, 16.45-18.00
Auditorium Bossa, Othon Palace Rio Hotel,
Av. Atlântica, 3264 - Copacabana,
Rio de Janeiro - RJ, 22070-001, Brazil.

http://www.win.tue.nl/ieeetfpm/
<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:45 – 16:50</td>
<td>Welcome</td>
</tr>
<tr>
<td>16:50 – 17:05</td>
<td>Overview of Activities and Developments in the Process Mining Space in 2015-2016 (Wil van der Aalst)</td>
</tr>
<tr>
<td></td>
<td><em>This session will provide a quick overview of the goals of the task force and its activities. It also reflects on its activities and the developments in the process mining space.</em></td>
</tr>
<tr>
<td>17:05 – 17:20</td>
<td>Process Mining: A Database of Applications by HSPI (Marcello La Rosa)</td>
</tr>
<tr>
<td></td>
<td><em>This session will present the results of a study aimed at building a comprehensive list of process mining applications. The study was conducted by HSPI, an Italian management consulting company. The session will briefly cover the method followed by the study and discuss the results. The list of applications will be distributed to the Task Force members afterwards.</em></td>
</tr>
<tr>
<td>17:20 – 17:40</td>
<td>Update and Overview of XES Standardization (Felix Mannhardt)</td>
</tr>
<tr>
<td></td>
<td><em>This session provides an overview on the XES Standardization efforts by the IEEE XES WG. The current state of the XES standard proposal is that it is under review by the IEEE. So far, we’re on track, and all lights are green. To get to this state, one issue needed to be resolved. Furthermore, one XES extension was approved. Both will be discussed, as will the action points for last year and data sets that have been published last year.</em></td>
</tr>
<tr>
<td>17:40 – 17:55</td>
<td>Discussion legal/ethical aspects of process mining (Andrea Burattin)</td>
</tr>
<tr>
<td></td>
<td><em>This session discusses the objectives and goals set for the &quot;privacy and ethical&quot; working group, with a status update for each of them after one year. Andrea will also present the documents/material collected so far, and some insights on the work done with an Italian lawyer.</em></td>
</tr>
<tr>
<td>17:55 – 18:00</td>
<td>Planned Activities for 2016-2017 (Wil van der Aalst)</td>
</tr>
<tr>
<td>18:00</td>
<td>Closing</td>
</tr>
<tr>
<td>Time</td>
<td>Session Title</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>16:45 – 16:50</td>
<td>Welcome</td>
</tr>
<tr>
<td>16:50 – 17:05</td>
<td><strong>Overview of Activities and Developments in the Process Mining Space in 2015-2016</strong> <em>(Wil van der Aalst)</em></td>
</tr>
<tr>
<td></td>
<td><em>This session will provide a quick overview of the goals of the task force and its activities. It also reflects on its activities and the developments in the process mining space.</em></td>
</tr>
<tr>
<td>17:05 – 17:20</td>
<td><strong>Process Mining: A Database of Applications by HSPI</strong> <em>(Marcello La Rosa)</em></td>
</tr>
<tr>
<td></td>
<td><em>This session will present the results of a study aimed at building a comprehensive list of process mining applications. The study was conducted by HSPI, an Italian management consulting company. The session will briefly cover the method followed by the study and discuss the results. The list of applications will be distributed to the Task Force members afterwards.</em></td>
</tr>
<tr>
<td>17:20 – 17:40</td>
<td><strong>Update and Overview of XES Standardization</strong> <em>(Felix Mannhardt)</em></td>
</tr>
<tr>
<td></td>
<td><em>This session provides an overview on the XES Standardization efforts by the IEEE XES WG. The current state of the XES standard proposal is that it is under review by the IEEE. So far, we’re on track, and all lights are green. To get to this state, one issue needed to be resolved. Furthermore, one XES extension was approved. Both will be discussed, as will the action points for last year and data sets that have been published last year.</em></td>
</tr>
<tr>
<td>17:40 – 17:55</td>
<td><strong>Discussion legal/ethical aspects of process mining</strong> <em>(Andrea Burattin)</em></td>
</tr>
<tr>
<td></td>
<td><em>This session discusses the objectives and goals set for the “privacy and ethical” working group, with a status update for each of them after one year. Andrea will also present the documents/material collected so far, and some insights on the work done with an Italian lawyer.</em></td>
</tr>
<tr>
<td>17:55 – 18:00</td>
<td><strong>Planned Activities for 2016-2017</strong> <em>(Wil van der Aalst)</em></td>
</tr>
<tr>
<td>18:00</td>
<td>Closing</td>
</tr>
</tbody>
</table>
Overview of Activities and Developments in the Process Mining Space in 2015-2016
Organizations Supporting our Task Force (70+)
PM Manifesto


Now in 16 Languages

New: Persian (Farsi), translated by Asef Pourmasoumi.

Opgoing discussion topic: Update of manifesto? Volunteers?
- Over 100,000 registrations
- Now in on-demand mode
- Updated based on new process mining book

https://www.coursera.org/learn/process-mining
Previous runs of the course (excl. on demand).

New on demand platform (course still being updated based on new book).
New MOOC:
Introduction to Process Mining with ProM

5697 registrations in first two runs
Next run: November 14th 2016

https://www.futurelearn.com/courses/process-mining
Other MOOCs?
Various PM Movies

IEEE CIS Task Force on Process Mining

Process Mining Movies

- Introduction to Process Mining: Turning (Big) Data into Real Value
- Process Mining In A Nutshell (Fluxicon)
- Perceptive Process Mining
- Automated Business Process Discovery - OPR ProcessAnalyzer
- Overview of the Business Process Insight (BPI) platform of IBM (IBM)
- Fuzzy model animation
- Process Mining animations using BPMiOne
- Interstage Automated Process Discovery demo
- Nomination of Process Mining Research at TU/e for ICT Regio Award
- Presentations on http://fluxicon.com/
- Prototyping Process Mining for SAP On
- Process mining 20sec
- Process mining
- Introduction to Process Mining
- ProcessMining - Improve Revenue, Rev
- Process Variants - Process Mining TV
- Process mining (Dutch, English subtitles)
- Perceptive Software & Alliander Price
- 15-min Video recording bpmNEXT press

Send links to your videos!

38.000+ views

Help to promote it: put it on your website!

https://www.youtube.com/watch?v=7oat7MatU_U
Examples of PM Movies

- [Celonis](https://www.youtube.com/watch?v=eUOutoXp75U)
- [QPR](https://www.youtube.com/watch?v=D0jaUMOU31I)
- [Perceptive](https://www.youtube.com/watch?v=nKy2Sx2WYRE)
- [Fluxicon](https://www.youtube.com/watch?v=KCpY90T3rQk)
LinkedIn Group on PM
(managed by George Varvaressos)

Now 5216 members

https://www.linkedin.com/grp/home?gid=1915049

1400 joined in the last year
Too little activity: Send more case studies!
Data Sets

Among the most viewed datasets on http://data.4tu.nl/
Irene Teinemaa, Anna Leontjeva and Karl-Oskar Masing. Diagnostics of Building Permit Application Process in Dutch Municipalities

Ube van der Ham. Benchmarking of Five Dutch Municipalities with Process Mining Techniques Reveals Opportunities for Improvement
Best Process Mining Dissertation Award 2015

Won by Jorge Munoz-Gama for his PhD thesis entitled “Conformance Checking and Diagnosis in Process Mining” (September 2015)

No award in 2016
More than 210 process mining practitioners from 165 different companies and 20 countries came together to learn from each other.
12th International Workshop on Business Process Intelligence 2016

to be held in conjunction with BPM 2016 Rio de Janeiro, Brasil, September 18 - 22, 2016

Business Process Intelligence Challenge (BPIC)

Sixth International Business Process Intelligence Challenge (BPIC'16)

In this challenge, we provide participants with a real-life event log, and we ask them to analyze these data using whatever techniques available, focusing on one or more of the process owner's questions or proving other unique insights into the process captured in the event log.

We strongly encourage people to use any tools, techniques, methods at their disposal. There is no need to restrict to open-source tools, and proprietary tools as well as techniques developed or implemented specifically for this challenge are welcome.

The winner will be announced at the Workshop on September 19th 2016.

Apply for a free license for process mining tool Minit v.2

GRADIENT ECM is providing free Minit licenses for participants of the Business Processing Intelligence Challenge 2016. Minit is an innovative software for automated discovery, analysis, auditing and optimization of business processes. Special licenses of the process mining tool include new functionality from version 2. To get expert answers to questions regarding the use of Minit please visit the BPIC Challenge 2016 Forum. Licenses are valid until 1st August 2016.

To apply for a free license please send an email to BPIC2016@gradientecm.com.

Important Dates

<table>
<thead>
<tr>
<th>Event</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract submission deadline</td>
<td>July 2016</td>
</tr>
<tr>
<td>Report submission deadline</td>
<td>July 2016</td>
</tr>
<tr>
<td>Announcement of winners</td>
<td>July 2016</td>
</tr>
</tbody>
</table>
Winners of the BPI Challenge 2016

DFKI/IWI team

Ube van der Ham of Meijer & Van der Ham Management Consultants
IEEE CIS Task Force on Process Mining

Process Discovery Contest @ BPM 2016

Background

Process Mining is a relatively young research discipline that sits between computational intelligence and data mining on the one hand, and process modeling and analysis on the other hand. The idea of process mining is to discover, monitor, and improve real processes (i.e., not assumed processes) by extracting knowledge from event logs readily available in today’s (information) systems. These techniques provide new means to discover, monitor, and improve processes in a variety of application domains. There are two main drivers for the growing interest in process mining. On the one hand, more and more events are being recorded, thus, providing detailed information about the history of processes. On the other hand, there is a need to improve and support business processes in competitive and rapidly changing environments. The lion’s share of attention of Process Mining has been devoted to Process Discovery, namely extracting process models - mainly business process models – from event logs.

The IEEE CIS Task Force on Process Mining aims to promote the research in the field of process mining and its application in real settings. In collaboration with it, to foster the research in the area of process discovery, we are proud to introduce the 1st Process-Discovery contest, which will be collocated with the BPM-2016 Conference in Rio de Janeiro in September 2016.

Objectives and Context

The Process Discovery Contest is dedicated to the assessment of tools and techniques that discover business process models from event logs. The objective is to compare the efficiency of process techniques to discover process models that provide a proper balance between “overfitting” and “underfitting”. A process model is overfitting (the event log) if it is too restrictive, disallowing behavior which is part of the underlying process. This typically occurs when the model only allows activities which are not part provided. These secret: only “training” technique that can avoid “underfitting”, that is, representing real executions of instances.

A model is as good in balancing “overfitting” and “underfitting” as it is able to correctly classify the traces in the “test” event log:
- Given a trace representing a behavior not related to the process, the model should classify it as allowed.
- Given a trace representing a behavior related to the process, the model should classify it as disallowed.

Josep Carmona, Universitat Politècnica de Catalunya
Massimiliano de Leoni, Eindhoven University of Technology
Benoît Depaire & Toon Jouck, Hasselt University
Winners of the PD Challenge 2016

Eric Verbeek and Felix Mannhardt (TU/e)
Process Miner of the Year Award

Joris Keizers
IEEE TSC special issue on "Processes meet Big Data"

Volume 8, Number 6, November - December 2015

- Wil M. P. van der Aalst, Ernesto Damiani: Processes Meet Big Data: Connecting Data Science with Process Science, 810-819
  [Electronic Edition BibTeX XML]
- Michael Werner, Nick Gehrke: Multilevel Process Mining for Financial Audits, 820-832
  [Electronic Edition BibTeX XML]
- Andrea Burattin, Marta Cimitle, Fabrizio Maria Maggi, Alessandro Sperduti: Online Discovery of Declarative Process Models from Event Streams, 833-846
  [Electronic Edition BibTeX XML]
  [Electronic Edition BibTeX XML]
- Xixi Lu, Marijn Nagelkerke, Dennis van de Wiel, Dirk Fahlend: Discovering Interacting Artifacts from ERP Systems, 861-873
  [Electronic Edition BibTeX XML]
  [Electronic Edition BibTeX XML]
2016 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)

Process Mining
Special Session at IEEE CIDM (SSCI) 2016

Home Call for Papers Organizers IEEE Task Force on Process Mining Process Mining Manifesto

28 days since the submission deadline

Submission Information
To submit your paper, please use the Manuscript System and select "Special Session for CIDM: Process Mining" as track. Please visit the main site for submission guidelines, including paper format and submission system.

Important dates
Paper submission: July 16-August 15
Decision: September 12
Final submission: October 10
Early registration: October 10

Andrea Burattin, University of Innsbruck, Austria
Fabrizio M. Maggi, University of Tartu, Estonia
Chiara Di Francescomarino, FBK, Italy

The IEEE Task Force on Process Mining is organizing a special session at the 2016 IEEE Symposium on Computational Intelligence and Data Mining (CIDM) 2016. The goal of this special session is to allow experts in the area of process mining and data analysis to share new techniques, applications and case studies. Therefore, submissions of papers on new process mining techniques, applications of process mining, business intelligence, process discovery, conformance checking, process intelligence, data analysis, etc. are welcome.

We now live in a time where the amount of data created daily goes easily beyond the storage and processing capabilities of nowadays systems. Organizations, governments but also individuals generate large amounts of data at a rate that has started to overwhelm the ability to timely extract useful knowledge from it. Nevertheless, the strategic importance of the knowledge hidden in these data is paramount for effective decision making and need to be extracted quickly in order to effectively react to dynamic situations. Efficient stream processing approaches for real-time analysis are crucial for enabling the predictive capabilities required by today's dynamically and rapidly evolving enterprises.

Process mining is a relatively young research discipline that sits between computational intelligence and data mining on the one hand and process modeling and analysis on the other. The idea of process mining is to discover, monitor and improve real processes (i.e., not assumed processes) by extracting knowledge from event logs readily available in today's systems.

Process mining provides an important bridge between data mining and business process analysis. Under the Business Intelligence (BI) umbrella many buzzwords have been introduced to refer to rather simple reporting and dashboard tools, such as BAN, CEP, CPM, CBI, BPI, TQM and Six Sigma. These approaches have in common that processes are just under a microscope to see whether further improvements are possible. Process mining is an enabling technology for BPM, BPI, TQM, Six Sigma, and the like.

Over the last decade, event data have become readily available and process mining techniques have matured. Process mining algorithms have been implemented in various academic and commercial systems. Today, there is an active group of researchers working on process mining and it has become one of the "hot topics" in Business Process Management (BPM) research. Moreover, there is a huge interest from industry in process mining. Many and more software vendors are adding process mining functionality to their tools.

Finally, the level of maturity and the relative low-cost of distributed approaches for storage and processing of information has not been fully exploited by the process mining community. There are very few research results on distributed storage methods and process mining algorithms.

Considering all these aspects, we believe the conference by enhancing possibilities. In addition it is a wonderful opportunity to meet everyone involved in the field.
XES standardization
Uptake of process mining

Documents by year

Scopus “process mining” papers 12-9-20176
<table>
<thead>
<tr>
<th>Short name</th>
<th>Full name of tool</th>
<th>Version</th>
<th>Vendor</th>
<th>Webpage</th>
<th>XES support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Celonis</td>
<td>Celonis Process Mining</td>
<td>4</td>
<td>Celonis GmbH</td>
<td><a href="http://www.celonis.de">www.celonis.de</a></td>
<td>yes</td>
</tr>
<tr>
<td>Disco</td>
<td>Disc</td>
<td>1.9.5</td>
<td>Fluxicon</td>
<td><a href="http://www.fluxicon.com">www.fluxicon.com</a></td>
<td>yes</td>
</tr>
<tr>
<td>EDS</td>
<td>Enterprise Discovery Suite</td>
<td>4</td>
<td>StereoLOGIC Ltd</td>
<td><a href="http://www.stereologic.com">www.stereologic.com</a></td>
<td></td>
</tr>
<tr>
<td>Fujitsu</td>
<td>Interstage Business Process Manager Analytics</td>
<td>12.2</td>
<td>Fujitsu Ltd</td>
<td><a href="http://www.fujitsu.com">www.fujitsu.com</a></td>
<td></td>
</tr>
<tr>
<td>Icaro</td>
<td>Icaro EVERFlow</td>
<td>1</td>
<td>Icaro Tech</td>
<td><a href="http://www.icarotech.com">www.icarotech.com</a></td>
<td></td>
</tr>
<tr>
<td>Icris</td>
<td>Icris Process Mining Factory</td>
<td>1</td>
<td>Icris</td>
<td><a href="http://www.processminingfactory.com">www.processminingfactory.com</a></td>
<td>yes</td>
</tr>
<tr>
<td>LANA</td>
<td>LANA Process Mining</td>
<td>1</td>
<td>Lana Labs</td>
<td><a href="http://www.lana-labs.com">www.lana-labs.com</a></td>
<td>not yet</td>
</tr>
<tr>
<td>Minit</td>
<td>Minit</td>
<td>1</td>
<td>Gradient ECM</td>
<td><a href="http://www.minitlabs.com">www.minitlabs.com</a></td>
<td>yes</td>
</tr>
<tr>
<td>myInvenio</td>
<td>myInvenio</td>
<td>1</td>
<td>Cognitive Technology</td>
<td><a href="http://www.my-invenio.com">www.my-invenio.com</a></td>
<td>soon (Nov)</td>
</tr>
<tr>
<td>Perceptive</td>
<td>Perceptive Process Mining</td>
<td>2.7</td>
<td>Lexmark</td>
<td><a href="http://www.lexmark.com">www.lexmark.com</a></td>
<td>no</td>
</tr>
<tr>
<td>ProM</td>
<td>ProM</td>
<td>6.6</td>
<td>Open Source hosted at TU/e</td>
<td><a href="http://www.promtools.org">www.promtools.org</a></td>
<td>yes</td>
</tr>
<tr>
<td>ProM Lite</td>
<td>ProM Lite</td>
<td>1.1</td>
<td>Open Source hosted at TU/e</td>
<td><a href="http://www.promtools.org">www.promtools.org</a></td>
<td>yes</td>
</tr>
<tr>
<td>QPR</td>
<td>QPR ProcessAnalyzer</td>
<td>2015.5</td>
<td>QPR</td>
<td><a href="http://www.qpr.com">www.qpr.com</a></td>
<td>yes</td>
</tr>
<tr>
<td>RapidProM</td>
<td>RapidProM</td>
<td>4.0.0</td>
<td>Open Source hosted at TU/e</td>
<td><a href="http://www.rapidprom.org">www.rapidprom.org</a></td>
<td>yes</td>
</tr>
<tr>
<td>Rialto</td>
<td>Rialto Process</td>
<td>1.5</td>
<td>Exeura</td>
<td><a href="http://www.exeura.eu">www.exeura.eu</a></td>
<td>yes</td>
</tr>
<tr>
<td>SNP</td>
<td>SNP Business Process Analysis</td>
<td>15.27</td>
<td>SNP Schneider-Neureither &amp; Partner AG</td>
<td><a href="http://www.snp-bpa.com">www.snp-bpa.com</a></td>
<td>yes</td>
</tr>
<tr>
<td>PPM</td>
<td>webMethods Process Performance Manager</td>
<td>9.9</td>
<td>Software AG</td>
<td><a href="http://www.softwareag.com">www.softwareag.com</a></td>
<td></td>
</tr>
</tbody>
</table>

Names missing? Send link.
New Vendors/tools
Growth: Celonis example

Celonis raises $27.5 million in funding from Accel Partners & 83North

Today Celonis, the leading provider of a new category of big data analytics technology dubbed Process Mining, announces a $27.5 million Series A investment led by Accel and 83North (formerly Greylock IL). Celonis will use the funds to accelerate the global adoption of its solution, which aims to help organisations of all sizes improve their operational efficiency by 20-30%.

Germany's fastest growing technology company

Celonis obtained the Deloitte Technology Fast 50 Award as the fastest growing technology company in Germany. With the Fast 50 Awards, Deloitte annually rewards the fastest growing companies in the high tech industry for their entrepreneurial excellence. Thanks to a growth rate of 3951 percent over the last four fiscal years, Celonis and its innovative and proprietary process mining software secured the first place. The award ceremony took place Monday night at a festive evening event at the Museum for Communication in Berlin. Prominent winners of the American Fast 50 Awards were, among others, Electric carmaker Tesla Motors and Google.

Process mining success story

Siemens

The world's largest process mining user
Anecdotes

Neighbor working in the hospital

Auditing of Tilburg university

Erik-Jan in train
Commercial versus Academic Tools

- Usability
- Speed
- Using additional data/dimensions
- Semantics or not?
- Conformance checking or filtering?
- Conformance checking or comparing graphs?

Convergence in coming years?
<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:45 – 16:50</td>
<td>Welcome</td>
</tr>
<tr>
<td>16:50 – 17:05</td>
<td><strong>Overview of Activities and Developments in the Process Mining Space in 2015-2016</strong> (Wil van der Aalst)</td>
</tr>
<tr>
<td>17:05 – 17:20</td>
<td><strong>Process Mining: A Database of Applications by HSPI</strong> (Marcello La Rosa)</td>
</tr>
<tr>
<td>17:20 – 17:40</td>
<td><strong>Update and Overview of XES Standardization</strong> (Felix Mannhardt)</td>
</tr>
<tr>
<td>17:40 – 17:55</td>
<td><strong>Discussion legal/ethical aspects of process mining</strong> (Andrea Burattin)</td>
</tr>
<tr>
<td>17:55 – 18:00</td>
<td><strong>Planned Activities for 2016-2017</strong> (Wil van der Aalst)</td>
</tr>
<tr>
<td>18:00</td>
<td>Closing</td>
</tr>
</tbody>
</table>
Research scope: geographical coverage

- Companies surveyed operate in various countries, with significant differences in terms of competitive scenarios, political, legal and tax context.
Process Mining Applications:
Database overview

Stefano Aiello, *partner*
Philippe Gonella, Serena Dal Molin, *consultants*
Vendor-independent management consulting company focusing on **IT Governance & Management**

- Funded in 2003, counts 100+ consultants
- Supports its customers in defining, designing, and implementing digital transformation strategies, enterprise architecture, and organizational changes, enabled through ICT
Research goals of Process Mining Applications Database

- Create a **comprehensive list** of process mining applications by collecting directly from the involved parties relevant information about the methodology used.

- Target audience: **executives, decision makers, companies** and **researchers** who want to know more about the business potentials of process mining.

✓ The ultimate goal of this knowledge endeavour is to build awareness and confidence about process mining applications in real life.
Research scope: geographical coverage

- Companies surveyed operate in various countries, with significant differences in terms of competitive scenarios, political, legal and tax context.
Six industries have been identified from the firms surveyed:

- Services
- Healthcare
- Utility
- Construction
- Manufacturing
- Chemical

The projects can be executed under the supervision or with the technical support of partners, such as universities, vendors, consultancy companies.

Timeline: from January 2005 to July 2016.
## A glimpse at the Database

<table>
<thead>
<tr>
<th>Category</th>
<th>Company/Agency</th>
<th>Process/Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing</td>
<td>Essmann Automotive (Germany)</td>
<td>By using process mining the, Essmann created efficient production processes. Its integration into lean corporate processes and into the management system, represented a key pillar of long-term success. By involving suppliers at an early stage in the product creation process and promoting team-oriented partnerships, Essmann achieved excellent and competitive products.</td>
</tr>
<tr>
<td>Service - Public</td>
<td>Centraal Bureau voor de Statistiek (Netherlands)</td>
<td>Statistics Netherlands is responsible for collecting and processing data in order to publish statistics to be used in practice, by policymakers and for scientific research. With process mining they have improved their performances.</td>
</tr>
<tr>
<td>Service - Credit</td>
<td>ING (Netherlands)</td>
<td>Making sure a customer has the best possible experience when interacting with the company is one the most important goals many companies strive for. ING DIRECT Australia asked for an in-depth analysis of the behaviour of their customers on their website before they called the call center. Using process mining they were able to get valuable business insights to make better decisions on how to further develop both their website and call center.</td>
</tr>
<tr>
<td>Service - Public</td>
<td>Ville de Lausanne (Switzerland)</td>
<td>Analysis of the construction permit process, in order to find bottlenecks.</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>Volvo (Germany)</td>
<td>A first dataset provided data about factory orders for the construction of trucks. The second dataset contained customer orders of trucks. It was discovered that the attribute 'ORDERNUMBER' of any event in a trace of the first log was also displayed in the attribute 'Omnnumber' of the event 'Accepted' in the second.</td>
</tr>
<tr>
<td>Healthcare</td>
<td>Scottish Rite Emergency Department of Children's Healthcare of Atlanta (US)</td>
<td>Process mining's visual analytics has played an important role in healthcare process analysis. The interactive visual approach enabled users to gain insight into the complexity of pediatric asthma care processes. It helped with care quality improvement programs, providing comparison, benchmarking and analysis of conformance to existing care protocols.</td>
</tr>
</tbody>
</table>

- 2014 - Celonis (Germany)  
- 2014 - Fluxicon (Netherlands)  
- 2014 - Ville de Lausanne (Switzerland)  
- 2014 - Ghent University, Department of Business Informatics and Operations Management, (Belgium)  
- 2014 - School of Interactive Computing & Tennenbaum Institute, Georgia Institute of Technology (US)
Process mining projects have significantly increased since 2005 (peak in 2014)

In just seven months in 2016, the cases already collected (14) are almost the same of those of the entire 2015 (15)
The most frequent applications are in the **service industry** (42%), followed by **healthcare** (39%) and **manufacturing** (14%).

Within service, **telco and public** (21% each), **insurance** (18%) and **banking** (10%).
The database shows that process enhancement (performance mining) is the process mining method applied most frequently (31%).

Nevertheless, very often a combination of methods is required (e.g. discovery + enhancement or conformance + enhancement).
According to the data, process mining is applied especially for improving the patient flow within a healthcare institution (25%), followed by IT service management processes (12%) and auditing processes (4%).

The remaining processes are industry-specific (e.g. claims handling processes).
Call for contributions!

- We would like to involve more organizations, academic institutions and experts in order to make this census even more useful, with the aim of creating the most comprehensive list of process mining applications worldwide.
- Such a source of knowledge can be built only through the collaboration of all the entities that are currently using, or that are willing to use in the next future, process mining techniques to improve their businesses.
- Therefore, we wish to keep in touch with anyone who is involved in real-life applications of process mining.

HSPI will release a new updated version of the database every year.
Call for Nominations 2017

- thesis **defended in 2015-2016**

- written in **English**, and of **sufficient quality**

- supervisors nominate until **1st May 2017**
  (see [http://tinyurl.com/pm-phd-award](http://tinyurl.com/pm-phd-award))

- monetary prize + LNBIP monography invitation
Section criteria

- Significance and innovation
- Technical depth:
  - conceptualization
  - formalization
  - implementation
  - evaluation
- Potential impact on academia and practice
- Quality of presentation
- Quality and number of publications
- Citations
Inaugural committee members (2014-2016):

• Antonella Guzzo (UniCal)
• Dirk Fahland (TU/e)
• Marcello La Rosa (QUT)

Anyone volunteering for 2017-2019?
<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Presenter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:45 – 16:50</td>
<td>Welcome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16:50 – 17:05</td>
<td>Overview of Activities and Developments in the Process Mining Space in 2015-2016</td>
<td>Wil van der Aalst</td>
<td><em>This session will provide a quick overview of the goals of the task force and its activities. It also reflects on its activities and the developments in the process mining space.</em></td>
</tr>
<tr>
<td>17:05 – 17:20</td>
<td>Process Mining: A Database of Applications by HSPI</td>
<td>Marcello La Rosa</td>
<td><em>This session will present the results of a study aimed at building a comprehensive list of process mining applications. The study was conducted by HSPI, an Italian management consulting company. The session will briefly cover the method followed by the study and discuss the results. The list of applications will be distributed to the Task Force members afterwards.</em></td>
</tr>
<tr>
<td>17:20 – 17:40</td>
<td>Update and Overview of XES Standardization</td>
<td>Felix Mannhardt</td>
<td><em>This session provides an overview on the XES Standardization efforts by the IEEE XES WG. The current state of the XES standard proposal is that it is under review by the IEEE. So far, we’re on track, and all lights are green. To get to this state, one issue needed to be resolved. Furthermore, one XES extension was approved. Both will be discussed, as will the action points for last year and data sets that have been published last year.</em></td>
</tr>
<tr>
<td>17:40 – 17:55</td>
<td>Discussion legal/ethical aspects of process mining</td>
<td>Andrea Burattin</td>
<td><em>This session discusses the objectives and goals set for the &quot;privacy and ethical&quot; working group, with a status update for each of them after one year. Andrea will also present the documents/material collected so far, and some insights on the work done with an Italian lawyer.</em></td>
</tr>
<tr>
<td>17:55 – 18:00</td>
<td>Planned Activities for 2016-2017</td>
<td>Wil van der Aalst</td>
<td></td>
</tr>
<tr>
<td>18.00</td>
<td>Closing</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
XES Standard

Felix Mannhardt, Eric Verbeek
Overview

Standard for XES – eXtensible Event Stream – for achieving interoperability in event logs and event streams.
Goal

• Facilitate export of event logs and event streams from an information system to an analysis tool
  – like, for example, a process mining tool
Scope

• To define a grammar for a language capturing event logs and event streams
  – XML Schema for event logs and event streams
  – XML Schema for extensions
  – Basic extensions
Working Group

- Wil van der Aalst
- Christian Günther
- Eric Verbeek
- J.C. Bose
- Josep Carmona
- Marlon Dumas
- Frank van Geffen
- Sukriti Goel
- Antonella Guzzo
- Rania Khalaf
- Rudolf Kuhn
- Teemu Lehto
- Felix Mannhardt
- Marco Montali
- Michael zur Muehlen
- Zbigniew Paszkiewicz
- Hajo Reijers
- Alexander Rinke
- Anne Rozinat
- Pnina Soffer
- Minseok Song
- Keith Swenson
- Walter Vanherle
- Lijie Wen
- Moe Wynn
Current Status

• Progress
  – XES Standard Proposal is on RevCom agenda
    • Sep 19 (that is, today)
  – Policies and Procedures submitted to sponsor
    • Pending
  – Micro extension approved

• We are on track!
XES Standard Proposal

- Dec 12: WG
- Jan 27: MEC
- Mar 18: Ballot
  - 1 issue
- Jul 7: Recirc
Balloting group

- Academic - Researcher: 32%
- Consulting: 18%
- General Interest: 23%
- Producer - System / Manufacturer: 23%
- User - Other: 4%
Issue: localtime

- xs:dateTime
  - This element doesn't specify "which time" is represented. It could be UTC, but also it could be "localtime" or "unspecified time".
  - I think standard should require all timestamps to be exact UTC, no timezone. And using "local time" should be banned - as "local time" is affected by change from/to DST; and order of events in such night would be e.g.: 1:00, 1:15, 1:30, 1:45, 1:00, 1:15; also these two events could be only 15 minutes apart: 1:00 and 2:15.
Micro extension

- Nested activities
  - Subactivities, micro activities, subprocess, ...
Micro extension proposal

- Start: Sep 17
- Comment: Nov 9
- Discuss: Dec 9
- Comment: Jan 4
- Approve: Jan 21
Action points

- Add XES support if not yet provided
  - QPR
- Add extensions and analysis techniques exploiting these
  - Micro
- Provide data sets via 3TU data center
  - Now called 4TU data center
  - ...
New data sets

November 3  Josep Carmona
• Activities of daily living of several individuals

April 22  Marcus Dees and Boudewijn van Dongen
• BPI Challenge 2016

July 5  Xixi Lu
• Handling Duplicated Tasks in Process Discovery by Refining Event Labels (BPM2016)

August 15  Maikel Leemans
• JUnit 4.12 Software Event Log
Looking ahead

- RevCom decides today
  - Recommendation to Standards Board

  “Upon approval by the IEEE-SA Standards Board, the standard shall be published as an IEEE standard. The Sponsor shall be notified of the approval in writing.”

- Update of OpenXES
- New extensions?
- New data sets?
<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:45 – 16:50</td>
<td>Welcome</td>
<td></td>
</tr>
<tr>
<td>16:50 – 17:05</td>
<td>Overview of Activities and Developments in the Process Mining Space in 2015-2016</td>
<td>Wil van der Aalst</td>
</tr>
<tr>
<td>17:05 – 17:20</td>
<td>Process Mining: A Database of Applications by HSPI</td>
<td>Marcello La Rosa</td>
</tr>
<tr>
<td>17:20 – 17:40</td>
<td>Update and Overview of XES Standardization</td>
<td>Felix Mannhardt</td>
</tr>
<tr>
<td>17:40 – 17:55</td>
<td>Discussion legal/ethical aspects of process mining</td>
<td>Andrea Burattin</td>
</tr>
<tr>
<td>17:55 – 18:00</td>
<td>Planned Activities for 2016-2017</td>
<td>Wil van der Aalst</td>
</tr>
<tr>
<td>18:00</td>
<td>Closing</td>
<td></td>
</tr>
</tbody>
</table>
Privacy in Process Mining

IEEE Task Force on Process Mining meeting
Agenda

• Organization of the group
• Goals set
• Update on goals
• Use case
• Discussion with a lawyer
• Bonus material
Members of the Working Group

• Members
  • Andrea Burattin, UIBK
  • Mauro Conti, UniPD
  • Felix Mannhardt, TU/e
  • Moe Wynn, QUT

• Anne Rozinat, Fluxicon
• Léonard Studer, Lausanne gov
• Daniele Turato, Siav
  • Roberto Pinelli, Siav (observer)

• Serena Faccio, Italian lawyer (interested)
Brief Recap of Main Events

• Asked some people for their interest (end September 2015)
• Mailing list setup (October 10th 201)
  • privacy_in_process_mining@listserv.tue.nl
• Some follow ups, but... everybody’s busy
Possible Goals (or Work Topics)

• **O1: Make a default (or example) NDAs**
  - Not to go into legal or country specific issues, but to show the typical ingredients and some examples
  - This is also related to publications and anonymization

• **O2: Define different levels of anonymization**
  - E.g., a ranking system to taxonomy of sensitivity of logs
  - Possibly include confidence (to decide how reliable are conclusions derived from it can be)

• **O3: Guidelines to remove sensitive information**
Regarding O1 (NDA)

- **Academic NDA by Moe Wynn, University of Queensland**
- It covers
  - Data access (identifiable vs non-identifiable data)
  - Publication consent (7/14 days)
  - Future use of data
  - Who can access the data – staff vs students (P/G and U/G)
  - Data privacy/storage
Regarding O1 (NDA)

• *Industrial* NDA, by Daniele Turato, SIAV

• It covers
  • Data access
  • Who can access the
  • Data privacy/storage

• In Italian only
Regarding O2/O3

• Conflicting goals: expressive power – privacy

• One idea could be to adopt $k$-anonymity
  • Aggregate information (e.g., instead of the exact activity time, just the day; instead of the resource name, the organizational unit)
  • Still, unsafe: no randomization, subjects may be identifiable. But is this enough?

• Differential privacy?
  • Idea: including data of one individual in the log is not strongly affecting the result so artificial noise makes impossible to know if someone participated
Use Case
Anonymizing a Hospital Event Log

• 1000 cases with potentially personally identifiable information
  • cases contain events that occurred for single patient (timestamps)
  • events contain additional attributes (diagnosis, age, test results)
  • event labels / terminology might hint at the data source

• Existing work/guidelines?
  • safe harbor guidelines
  • data sets used in health care (medical studies), and many more?

• How to anonymize for Process Mining?
  • most existing research does not apply directly to event log
  • high risk of re-identification due to combined of timestamps & attributes in event logs
  • balance between usefulness/realism and anonymization
Used methods available in ProM

- **Remove attributes**
  - Examples: Remove technical identifiers, literal names, ...

- **Rename events**
  - Example: Replace “Release to home care” with “Release A”, ...

- **Rename literal attributes**
  - Example: Replace attribute “Diagnosis Pneumonie” with “Diagnosis A”

- **Anonymize numeric attributes**
  - Example: Generalize “31 yr” with “Age 30-40”

- **Anonymize timestamps**
  - Example: Move all timestamps of a trace by a random amount (Caution: changes arrival rate, etc.)

- **Package: LogEnhancement**
Resulting Event Log

- 1000 cases without (hopefully) personally identifiable information
  - One case: a patient visited the ER room with a suspicion for SEPSIS
  - 16 event classes, 31 attributes
- Replaces event labels with generic labels
  - Avoids that people with insider information can recognize the data
  - Avoids that the data can be combined with other data
  - Side effect: the labels are easier to understand for others
- Generalized age into age groups (31 -> 30-35)
  - One group for >90 to avoid groups with small population
- Generalized specific diagnosis into generic labels (A,B,C,...)
  - Limits the usefulness somewhat, but fine to test algorithms
- Renamed some attributes with generic names (CheckboxA, CheckboxB, ...)
- Randomly shifted the time stamps of traces
  - Relative time between events kept
  - Prevents any kind of analysis wrt. inter trace metrics (arrival time, etc.)
Discussion with a Lawyer
Process Mining and Privacy: Roles

Data collection
(company/institute side)

Data and process owners are in charge of legal data collection.

If there is no authorization for data collection, then nothing can be done.

Data/process mining analysis

Data analysis should just be based on what the data and process owners allowed during the data collection process. The rest is not allowed.

(Privacy enhanced techniques are not strongly needed)
Data Collection

• Process mining deals with employees data
  • With clients data easier: the contract might include *data analysis*

• Data collection has to be legitimate
  • Employee work is protected by privacy regulations
  • Privacy is recognized as a fundamental right
    • Article 8 of *European Convention on Human Rights*
    • Article 12 of *Universal Declaration of Human Rights*
    • It is not possible to delegate the data collection consent

• Data collection consent
  • Cannot be forced (reward-based approach instead?)
  • Has to allow the analysis involved
Data Processing

• New EU regulation
  • Approved on April 27th 2016
  • Shall apply from May 25th 2018

• REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL
  on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

The New Regulation

Article 4 - Definitions

For the purposes of this Regulation:

• (1) ‘personal data’ means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;

• (2) ‘processing’ means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction;

• (4) ‘profiling’ means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements;
The New Regulation

Article 6 - Lawfulness of processing

1. Processing shall be lawful only if and to the extent that at least one of the following applies:

• (a) the data subject has given consent to the processing of his or her personal data for one or more specific purposes;

• (b) processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract

• ...
The New Regulation

Article 9 - Processing of special categories of personal data

1. Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited.

2. Paragraph 1 shall not apply if one of the following applies:

• (a) the data subject has given explicit consent to the processing of those personal data for one or more specified purposes, except where Union or Member State law provide that the prohibition referred to in paragraph 1 may not be lifted by the data subject;

• (b) processing is necessary for the purposes of carrying out the obligations and exercising specific rights of the controller or of the data subject in the field of employment and social security and social protection law in so far as it is authorised by Union or Member State law or a collective agreement pursuant to Member State law providing for appropriate safeguards for the fundamental rights and the interests of the data subject;
The New Regulation

**Article 88 - Processing in the context of employment**

1. Member States may, by law or by collective agreements, **provide for more specific rules** to ensure the protection of the rights and freedoms in respect of the processing of employees' personal data in the employment context, in particular for the purposes of the recruitment, the **performance of the contract of employment**, including discharge of obligations laid down by law or by collective agreements, management, planning and organisation of work, equality and diversity in the workplace, health and safety at work, protection of employer's or customer's property and for the purposes of the exercise and enjoyment, on an individual or collective basis, of rights and benefits related to employment, and for the purpose of the termination of the employment relationship.
Open Problems in Law Perspective

• Interpretation of points (b) of art. 6 and 9 Reg.

• Transitivity of consent to processing data for a new process (not foreseen before but related or necessary for the analysis)

• To what extent EU countries can override the general principles in order “to ensure the protection of the rights and freedoms in respect of the processing of employees' personal data in the employment context” (art. 88 Reg.)
<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:45 – 16:50</td>
<td>Welcome</td>
</tr>
</tbody>
</table>
| 16:50 – 17:05 | Overview of Activities and Developments in the Process Mining Space in 2015-2016  
(Wil van der Aalst) |
| 16:50 – 17:05 | This session will provide a quick overview of the goals of the task force and its activities. It also reflects on its activities and the developments in the process mining space. |
| 17:05 – 17:20 | Process Mining: A Database of Applications by HSPI  
(Marcello La Rosa) |
| 17:05 – 17:20 | This session will present the results of a study aimed at building a comprehensive list of process mining applications. The study was conducted by HSPI, an Italian management consulting company. The session will briefly cover the method followed by the study and discuss the results. The list of applications will be distributed to the Task Force members afterwards. |
| 17:20 – 17:40 | Update and Overview of XES Standardization  
(Felix Mannhardt) |
| 17:20 – 17:40 | This session provides an overview on the XES Standardization efforts by the IEEE XES WG. The current state of the XES standard proposal is that it is under review by the IEEE. So far, we’re on track, and all lights are green. To get to this state, one issue needed to be resolved. Furthermore, one XES extension was approved. Both will be discussed, as will the action points for last year and data sets that have been published last year. |
| 17:40 – 17:55 | Discussion legal/ethical aspects of process mining  
(Andrea Burattin) |
| 17:40 – 17:55 | This session discusses the objectives and goals set for the "privacy and ethical" working group, with a status update for each of them after one year. Andrea will also present the documents/material collected so far, and some insights on the work done with an Italian lawyer. |
| 17:55 – 18:00 | Planned Activities for 2016-2017  
(Wil van der Aalst) |
| 17:55 – 18:00 | This session will present the results of a study aimed at building a comprehensive list of process mining applications. The study was conducted by HSPI, an Italian management consulting company. The session will briefly cover the method followed by the study and discuss the results. The list of applications will be distributed to the Task Force members afterwards. |
| 18.00 | Closing |

18.00
Planned Activities for 2016-2017
Continued Activities
(assumption)

• Co-located with BPM 2017 in Barcelona:
  – BPI Workshop
  – BPI Challenge
  – Process Mining Contest
  – Best Process Mining Dissertation Award
• Algorithms & Theories for the Analysis of Event Data (ATAED 2017, Zaragoza, Spain)
• Process Mining Camp (Fluxicon), Eindhoven
• XES standardization
• …
Volunteers needed!

• Process Mining Session at the IEEE World Congress on Computational Intelligence (WCCI 2018), Rio de Janeiro, Brazil, July 2018.


• Dagstuhl Seminar on PM? (http://www.dagstuhl.de/en/program/dagstuhl-seminars/proposal/ deadline November 1, 2016)

• Special issues of Computing, BISE, …?

Calendar (volunteer)? Geographic and temporal distribution?
Input Needed

• Additional Use Cases
• Additional Public Data Sets
• Additional MOOCs, books, etc. for end-users
Any new ideas?

Use the task force!

W: www.vdaalst.com
T: @wvdaalst
E: w.m.p.v.d.aalst@tue.nl
Closing

Thanks!!

Next: Welcome Reception