XES Standard

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
XES Standard Proposal

- Recirc: Jul 7
- Approved: Sep 26 • D03
- Draft: Sep 30
- Published: Nov 11 • 1849-2016
Abstract:
A grammar for a tag-based language whose aim is to provide designers of information systems with a unified and extensible methodology for capturing systems behaviors by means of event logs and event streams is defined in the XES standard. An XML Schema describing the structure of an XES event log/stream and a XML Schema describing the structure of an extension of such a log/stream are included in this standard. Moreover, a basic collection of so-called XES extension prototypes that provide semantics to certain attributes as recorded in the event log/stream is included in this standard.

Abstract:

Presents information on the development of IEEE Standard 1849, the XES Standard.

Published in: IEEE Computational Intelligence Magazine (Volume: 12, Issue: 2, May 2017)
XES Certification Proposal

- **RfC1** • Jun 7
- **RfC2** • June 21
- **RfA** • Jul 5
- **Approved** • Jul 19
XES Certification
SW Event  Ext. Proposal

- RfC1 • Jul 20
- RfC2 • Aug 24
- RfA • Sep 7
- Now • Sep 21
- Approved? • Sep 21
SW Event Extension

- Software events
Action points

• XES Certification
  – Certify existing tools

• XES Software Extensions
  – Software Event (to be approved)
  – Software Communication (to be proposed)
  – Software Telemetry (to be proposed)

• Provide data sets via 4TU data center
New data sets

December 7  Felix Mannhardt
• Sepsis Cases- Event Log

February 6  Boudewijn van Dongen
• BPI Challenge 2017

April 21  Maikel Leemans
• NASA Crew Exploration Vehicle (CEV) Software Event Log

April 21  Maikel Leemans
• Apache Commons Crypto 1.0.0 - Stream CbcNopad Unit Test Software Event Log
XES Certification

Proposal on how to obtain a certification for your tool
Tools supporting XES

- Celonis Process Mining
- Disco
- Icris Process Mining Factory
- Minit
- myInvenio
- ProcessGold Enterprise Platform
- ProM
- ProM Lite
- QPR Process Analyzer
- RapidProM
- Rialto Process
- SNP Business Process Analysis
- ...
Motivation

• What does it mean if two different tools support XES?
• Can one tool successfully import a XES log file that was exported by another tool?
Levels
Certification Objectives

**Import**
- Show an artifact, for example a discovered process model, that shows that the relevant imported attributes have been used.

**Export**
- Export the data, as loaded in the tool, to a XES event log that shows that the relevant loaded attributes have been exported.

- **B1:**
  - Tag
  - Recycle
  - Clock

- **C2:**
  - Magnifying Glass
  - User
Import: Logs

Real-life

• Sanitized BPIC logs
• Tools should not crash when importing these log files

Artificial

• Tools should not crash, but should furthermore show that the relevant attributes from every log were used
# Real-life logs

<table>
<thead>
<tr>
<th>Event log</th>
<th>Traces</th>
<th>Events</th>
<th>Size in KB</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPIC12</td>
<td>13,087</td>
<td>262,200</td>
<td>72,363</td>
</tr>
<tr>
<td>BPIC13_closed_problems</td>
<td>1,487</td>
<td>6,660</td>
<td>4,090</td>
</tr>
<tr>
<td>BPIC13_incidents</td>
<td>7,554</td>
<td>65,533</td>
<td>38,627</td>
</tr>
<tr>
<td>BPIC13_open_problems</td>
<td>819</td>
<td>2,351</td>
<td>1,370</td>
</tr>
<tr>
<td>BPIC15_1</td>
<td>1,199</td>
<td>52,217</td>
<td>40,261</td>
</tr>
<tr>
<td>BPIC15_2</td>
<td>832</td>
<td>44,354</td>
<td>33,616</td>
</tr>
<tr>
<td>BPIC15_3</td>
<td>1,409</td>
<td>59,681</td>
<td>45,673</td>
</tr>
<tr>
<td>BPIC15_4</td>
<td>1,053</td>
<td>47,293</td>
<td>36,131</td>
</tr>
<tr>
<td>BPIC15_5</td>
<td>1,156</td>
<td>59,083</td>
<td>44,961</td>
</tr>
<tr>
<td>BPIC17 – Offer log</td>
<td>42,995</td>
<td>193,849</td>
<td>107,557</td>
</tr>
<tr>
<td>BPIC17</td>
<td>31,509</td>
<td>1,202,267</td>
<td>565,373</td>
</tr>
</tbody>
</table>
## Artificial logs

<table>
<thead>
<tr>
<th>Event log</th>
<th>Level / Flag</th>
<th>Event attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>LevelA1</td>
<td>A1</td>
<td>🔄</td>
</tr>
<tr>
<td>LevelA2</td>
<td>A2</td>
<td>🔄 (➕ + 🔄)</td>
</tr>
<tr>
<td>LevelB1</td>
<td>B1</td>
<td>🔄, 🔄, 🟥</td>
</tr>
<tr>
<td>LevelB2</td>
<td>B2</td>
<td>🔄 (➕ + 🔄), 🟥</td>
</tr>
<tr>
<td>LevelC1</td>
<td>C1</td>
<td>🔄, 🟥</td>
</tr>
<tr>
<td>LevelC2</td>
<td>C2</td>
<td>🔄 (➕ + 🔄), 🟥</td>
</tr>
<tr>
<td>LevelD1</td>
<td>D1</td>
<td>🔄, 🖢, 🔄, 🟥, 🗻, 🎓, 🕒</td>
</tr>
<tr>
<td>LevelD2</td>
<td>D2</td>
<td>🔄 (➕ + 🔄), 🖢, 🟥, 🗻, 🎓, 🕒</td>
</tr>
<tr>
<td>FlagX1</td>
<td>X1</td>
<td>🤕 (only non-standard attributes)</td>
</tr>
<tr>
<td>FlagX2</td>
<td>X2</td>
<td>🔄, 🚨 (only non-standard attributes)</td>
</tr>
</tbody>
</table>
Process

1. Report, by vendor
   - Template

2. Demo, by vendor and XES WG rep
   - Preferably live, screencast acceptable
   - In case of issues: back to Report

3. Approval, by XES WG
   - Within 2 weeks

4. Publication, by XES WG rep
   - On the TF website
<table>
<thead>
<tr>
<th>Tool</th>
<th>Meta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Authors</td>
</tr>
<tr>
<td>Vendor</td>
<td>Date</td>
</tr>
<tr>
<td>Version</td>
<td>History</td>
</tr>
<tr>
<td>Requested cert. level(s)</td>
<td>Import (if applicable)</td>
</tr>
<tr>
<td></td>
<td>Export (if applicable)</td>
</tr>
</tbody>
</table>
Report (2)

Import

- Name of import log
- Walk-through with screenshots showing that the log is imported successfully
- If artificial log: Walk-through with screenshots showing that the attributes contained in the log were used

Export

- Walk-through with screenshots showing that the current data as stored by the tool is exported successfully to a log
- Walk-through with screenshots showing that the attributes as contained in the log match the current data as stored by the tool
Discussion

Comments?

Approval?