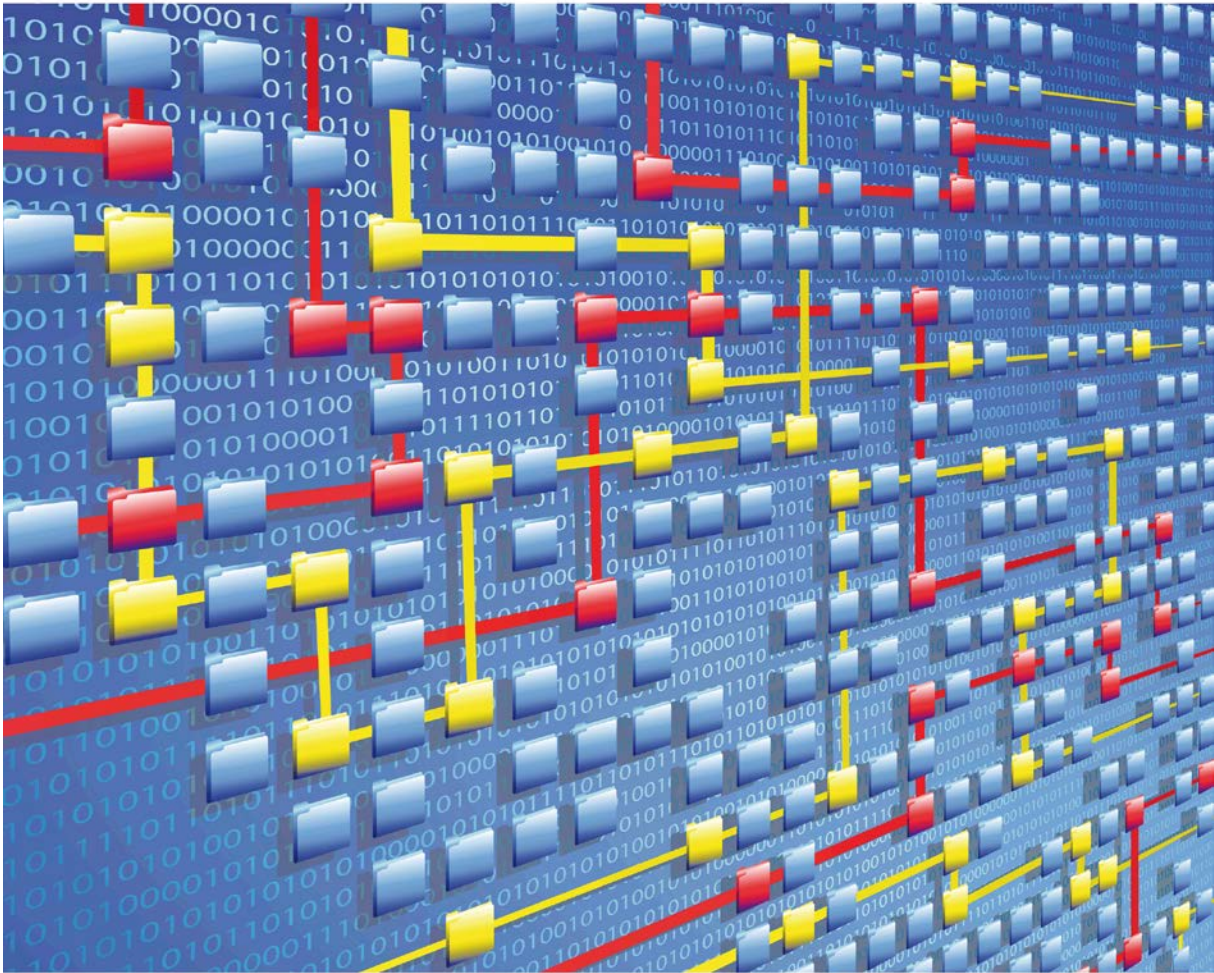


IEEE Task Force on



Process Mining

XES CERTIFICATION

XES WG 2017

TABLE OF CONTENTS

Contents

Introduction	1
Certification Levels	2
Successful Export/Import Combinations	4
Example Certificates	5

Introduction

XES STANDARD

On September 22nd, 2016, the **NEW P1849 (CIS/SC) IEEE Draft Standard for XES - eXtensible Event Stream - For Achieving Interoperability in Event Logs and Event Streams** was approved by the IEEE-SA Standards Board. On November 11th, 2016, IEEE Standards Association has officially published the XES Standard as **IEEE Std 1849TM-2016: IEEE Standard for eXtensible Event Stream (XES) for Achieving Interopability in Event Logs and Event Streams**. The IEEE Task Force on Process Mining has been driving the standardization process for over six years, because the standard allows for the exchange of event data between different process mining tools.

Through the XES Standard, event data can be transported from the location where it was generated to the location where it can be stored and analyzed, without losing semantics. The XES Standard enforces that this transport and storage is done in a standardized way, that is, in a way that is clear and well-understood. Next to providing a standardized syntax and semantics, the XES Standard also allows for extensions, e.g., adding cost information or domain specific attributes to events.

XES CERTIFICATION

Even though different tools may use the XES Standard for importing and exporting event data, there is no guarantee that the export of one tool can be successfully imported in another tool. Whether this import can be done, depends on whether the actual content of the XES event data matches the assumptions of the import.

To facilitate achieving interoperability between different tools on the XES Standard, we propose a XES Certificate. This Certificate assigns multiple XES certification levels to a tool, both of the import and the export. Based on these certification levels, it is straightforward to check that the export of one tool can indeed be successfully imported in another tool.

DOCUMENT STRUCTURE

The remainder of this document is structured as follows. The next section introduces the four different certification levels, from A to D together with an out-of-level flag X. The third section then shows an overview of which combinations of exports and imports lead to a successful import. Finally, the fourth section shows an example overview of certification levels for some of the tools. Of course, this overview is still premature and needs to be corrected, extended, and completed with other XES-aware tools in the future. The overview just serves as an example.

Wil van der Aalst, Christian Günther, and Eric Verbeek
May 31, 2017

CERTIFICATION LEVELS

Certification Levels

We distinguish four levels of XES event data, from A to D:

- A. XES Event data that only contains case identifiers and activity names. This level allows for process discovery and conformance checking.
- B. Level A extended with event types (lifecycle:transition attribute) and timestamps (time:timestamp attribute). In addition, this level allows for performance analysis, and for the generation of basic simulation models.
- C. Level B extended with information on resources (org:resource attribute). In addition, this level allows for social network mining, and for the generation of more advanced (resource-aware) simulation models.
- D. Level C extended with attributes from any standard XES extension. As a result, event data on this level still has clear semantics.

Beyond these four levels, we can have XES event logs that contain attributes from non-standard XES extensions and/or attributes without an extension. For these logs, we introduce the out-of-level flag X:

- X. Level D extended with attributes from non-standard XES extensions and/or attributes without an extension. As non-standard XES extension may conflict on the semantics of certain attributes, and the semantics of extension-less attributes is not clear, the semantics of event data flagged with X is possibly not clear.

For the Level A support, we foresee two possibilities: 1 and 2:

1. Fixed support, meaning that for both the case identifiers and the activity names the concept:name attribute is used.
2. Flexible support, meaning that for the case identifiers a trace classifier is used, and for the activity names an event classifier is used. Note that the event data can contain multiple trace classifiers and multiple event classifiers. In such cases, by default the first classifiers are selected, or (if the tool permits) other classifiers may be selected by the user.

A tool that only supports possibility 1 will be certified as Level A1, a tool that only supports possibility 2 as Level A2, and a tool that supports both as Level A. As Level A underlies all other levels, we also make this distinction on the other levels. For example, a tool that supports Level B but only Level A1 will be certified as B1, a tool that supports Level B but only Level A2 as B2, and a tool that supports B and both A1 and A2 as simply B.

Finally, we distinguish export from import, as the levels on both may not be the same. A tool that has Level B for export can write XES event data that includes:

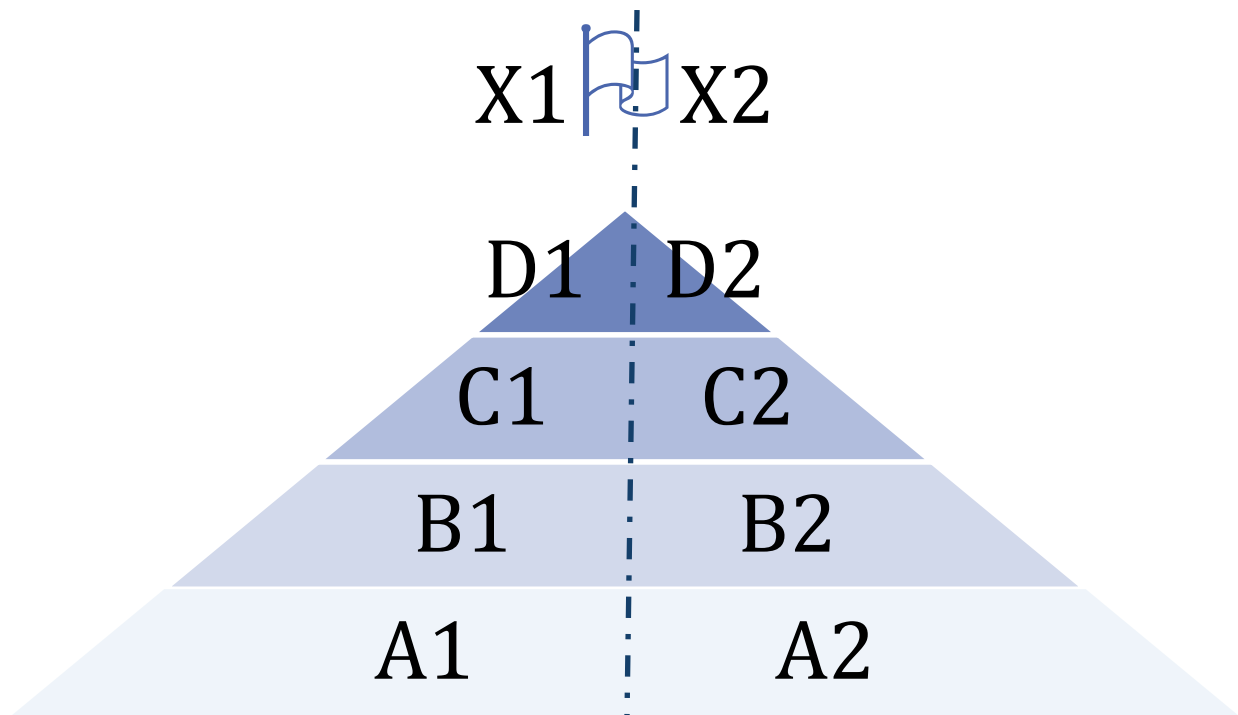
- case identifiers and activity names using either the fixed or the flexible support, and
- event types and timestamps.

CERTIFICATION LEVELS

A tool that has Level B1 for import can successfully (in a semantically correct way) read XES event data that includes:

- case identifiers and activity names using the fixed but not the flexible support, and
- event types and timestamps.

The following figure shows the different levels and the out-of-level flag in a graphical way (where the union of A1 and A2 yields A, etc.):



SUCCESSFUL EXPORT/IMPORT COMBINATIONS

Successful Export/Import Combinations

The following table shows the export/import combinations that are successful (+), not successful (-), or where the success depends on undefined attributes (?). Note that for the successful combinations we assume that there is a match on Level A, that is, both tools support the fixed and/or the flexible possibility. If not, that is, if one tool only supports Level A1 while the other only supports Level A2, then the import will not be successful.

EXPORT/IMPORT MATRIX

Import	Level A	Level B	Level C	Level D	Flag X
<i>Export Level A</i>	+	-	-	-	-
<i>Level B</i>	+	+	-	-	-
<i>Level C</i>	+	+	+	-	-
<i>Level D</i>	+	+	+	+	-
<i>Flag X</i>	+	+	+	+	?

As the table shows, a tool that requires an X-flagged event log for import has its problems with importing XES event data, as it cannot be sure that it will interpret all event data in a correct way. Basically, the owners of such tools should strive to have the non-standardized attributes covered by a (new) standard XES extension, which would automatically move the import to Level D.

EXAMPLE CERTIFICATES

Example Certificates

The following table shows example Certificates (with the certified levels) for some of the tools that support XES. Note that we do not claim that this overview is complete nor that this list is necessarily correct. As mentioned in the introduction, this overview is premature, and needs to be corrected, extended, and completed with other XES-aware tools. Note that a proper overview can only be made if the Certification has been fixed, and that the latter is exactly the aim of the current proposal.

EXAMPLE CERTIFICATES		
TOOL	IMPORT LEVEL	EXPORT LEVEL
Celonis PI	B1	-
Disco	A1, B1	D1
ProM 6	A-D, X	A-D, X

As an example, this table shows that we can successfully transfer the event data as loaded in Disco to Celonis PI, but not the other way around. As another example, this table shows that Celonis PI and Disco cannot handle classifiers, while ProM 6 can. As a third example, this table shows that Celonis PI requires Level B1 for the import, as it requires the timestamps to be included in the event data. As such, it cannot successfully import Level A1 event data.