

Dataset **CDISC StudyDataSet-XML**

Leaving the Stone Age of data transmission

PhUSE SDE Copenhagen, 11th June 2014
Sven Greiner
Statistical Programming, Accovion GmbH

Contents

- I. Dataset-XML – what and why?
- II. Introduction to XML & ODM
- III. Implementing Dataset-XML
- IV. Dataset-XML Tools
- V. Next Steps

I. Dataset-XML – what and why?

What is Dataset-XML?

- ✓ Defines format for transporting datasets in XML
- ✓ Based on the Operational Data Model (ODM)
- ✓ Supports ADaM, SDTM, SEND and other data
- ✓ Transport of datasets in FDA submissions

I. Dataset-XML – what and why?

Why Dataset-XML?

- ✓ FDA recommends SAS Transport v5 in 1999
- ✓ Limitations:
 - 8 char variable names, 200 char variable length, 40 char label length
 - Huge dataset sizes
 - ...
- ✓ FDA in November 2012: Dataset-XML an alternative for consideration



II. Introduction to XML & ODM

The Extensible Markup Language (XML)

- ✓ Open standard produced by the W3C
- ✓ XML is a textual data format
- ✓ Data has to conform to an XML schema

```
<svg xmlns="http://www.w3.org/2000/svg" version="1.1" width="500"
height="400">
```

```
<rect x="0" y="210" width="300" height="240" fill="blue" />
```

```
<ellipse cx="280" cy="230" rx="190" ry="120" fill="yellow"/>
```

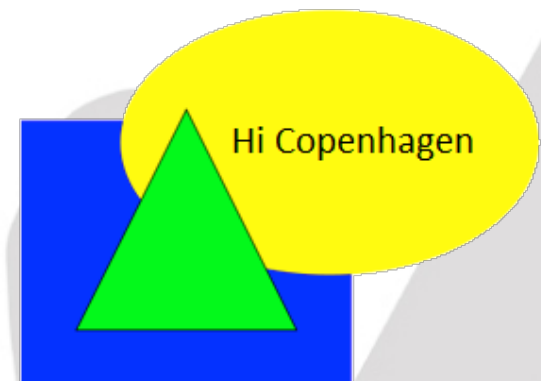
```
<path d="M150 200 L50 400 L250 400 Z" stroke="black" fill="lime" />
```

```
<text x="180" y="240" font-family="Arial" font-size="30">
```

```
  Hi Copenhagen
```

```
</text>
```

```
</svg>
```



II. Introduction to XML & ODM

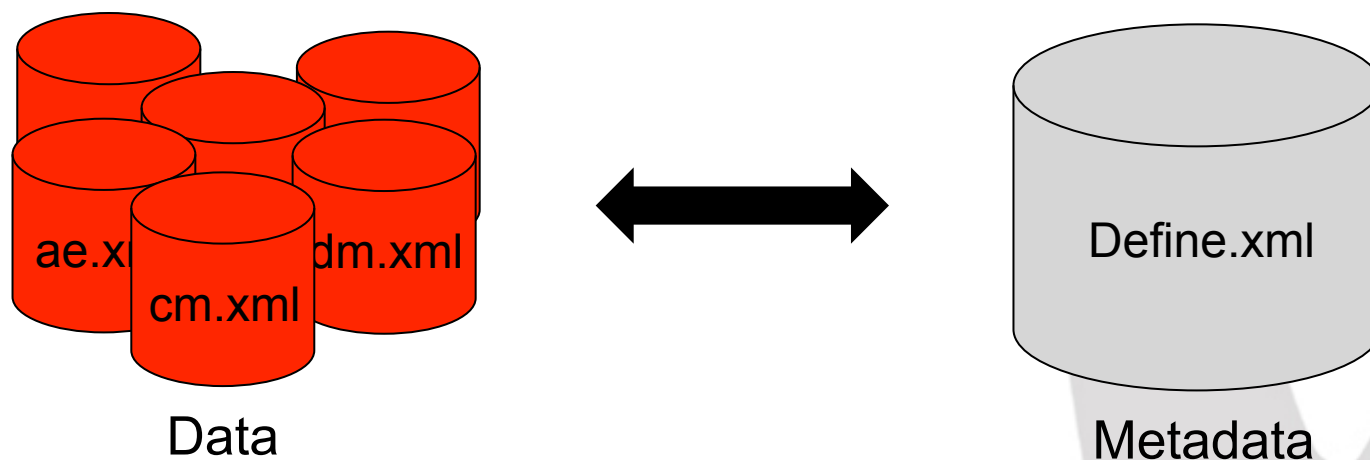
The Operational Data Model (ODM)

- ✓ Format for the interchange and archival of clinical study data using XML
- ✓ Includes:
 - Clinical data, associated metadata, administrative data, reference data and audit information
- ✓ Covers all aspects of clinical research data

III. Implementing Dataset-XML

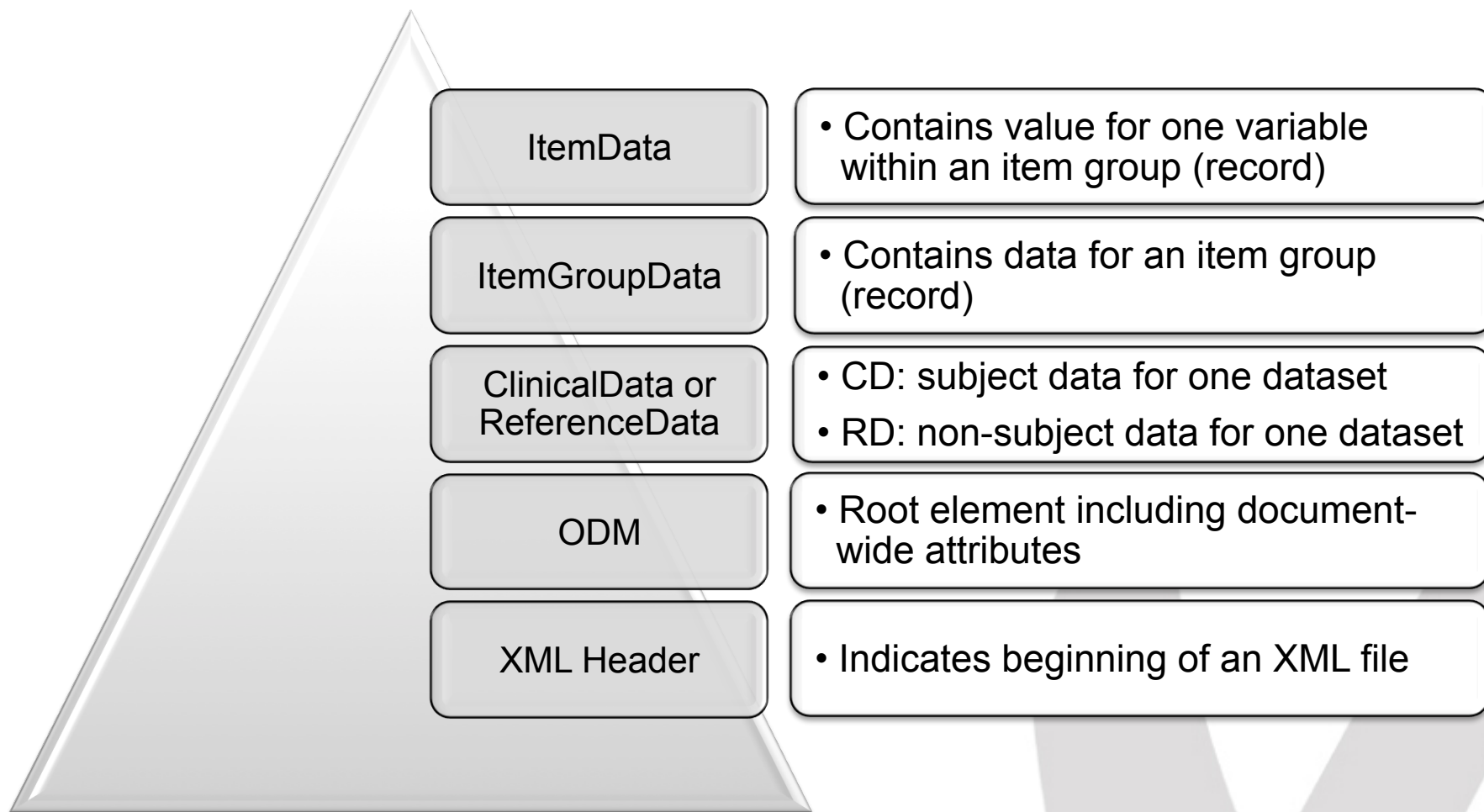
Features of Dataset-XML

- ✓ Dataset-XML is an extension of ODM
- ✓ One XML-file per dataset
- ✓ Metadata stored outside the dataset (Define.xml)



III. Implementing Dataset-XML

Dataset-XML elements



III. Implementing Dataset-XML

CM example



Cm.xml

The example file is part of the „CDISC Dataset-XML Specification Version 1.0“ package.

IV. Dataset-XML Tools

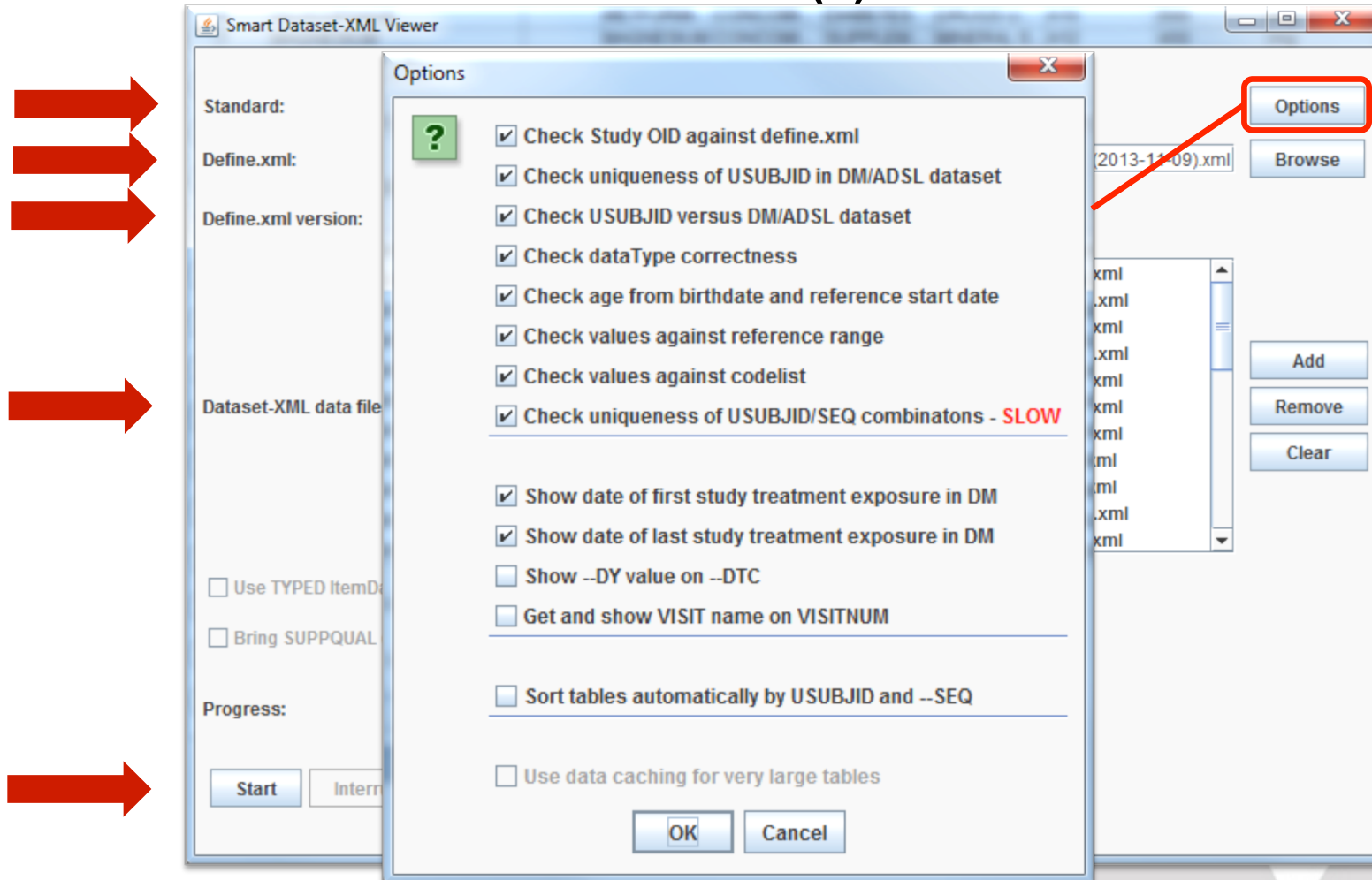
Overview

Tool	Description
EZ Convert	✓ Converts Dataset-XML files into SAS datasets
SAS Clinical Standards Toolkit	✓ Dataset-XML support will be part of the next release of CST
OpenCDISC v1.5	✓ OpenCDISC v1.5 works with Dataset-XML files and Define-XML v2.0
XPT2DatasetXML	✓ Transforms XPT datasets into Dataset-XML datasets
Smart Dataset-XML Viewer	✓ Shows Dataset-XML files as tabular datasets

Source: <http://wiki.cdisc.org/display/PUB/CDISC+Dataset-XML+Resources>

IV. Dataset-XML Tools

Smart Dataset-XML Viewer (1)



V. Next Steps

What is next for Dataset-XML?

✓ FDA

- Complete the pilot project (January 2015?)
- Create infrastructure
- Allow Dataset-XML for submissions

✓ Industry

- Wait for FDA decision
- Staff training
- Adjust processes



Questions?

ACCOVION

Sven Greiner

Senior Statistical
Programmer

www.accovion.com

Accovion GmbH
Helfmann-Park 10
D-65760 Eschborn, Germany
Tel. +49 6196 7709-288
sven.greiner@accovion.com