

iiWAS 2004 Invited Speach

**Advanced Cross-Publishing with XML
and XSL using Platform Independent
Open Source Tools**

Alexander Schatten
Vienna University of Technology

Content

1. Cross Publishing and System Integration Introduction

- Cross-Publishing / Multi-Channel Publishing
- System Integration
- Sources and Targets
- Open Source/Protocol

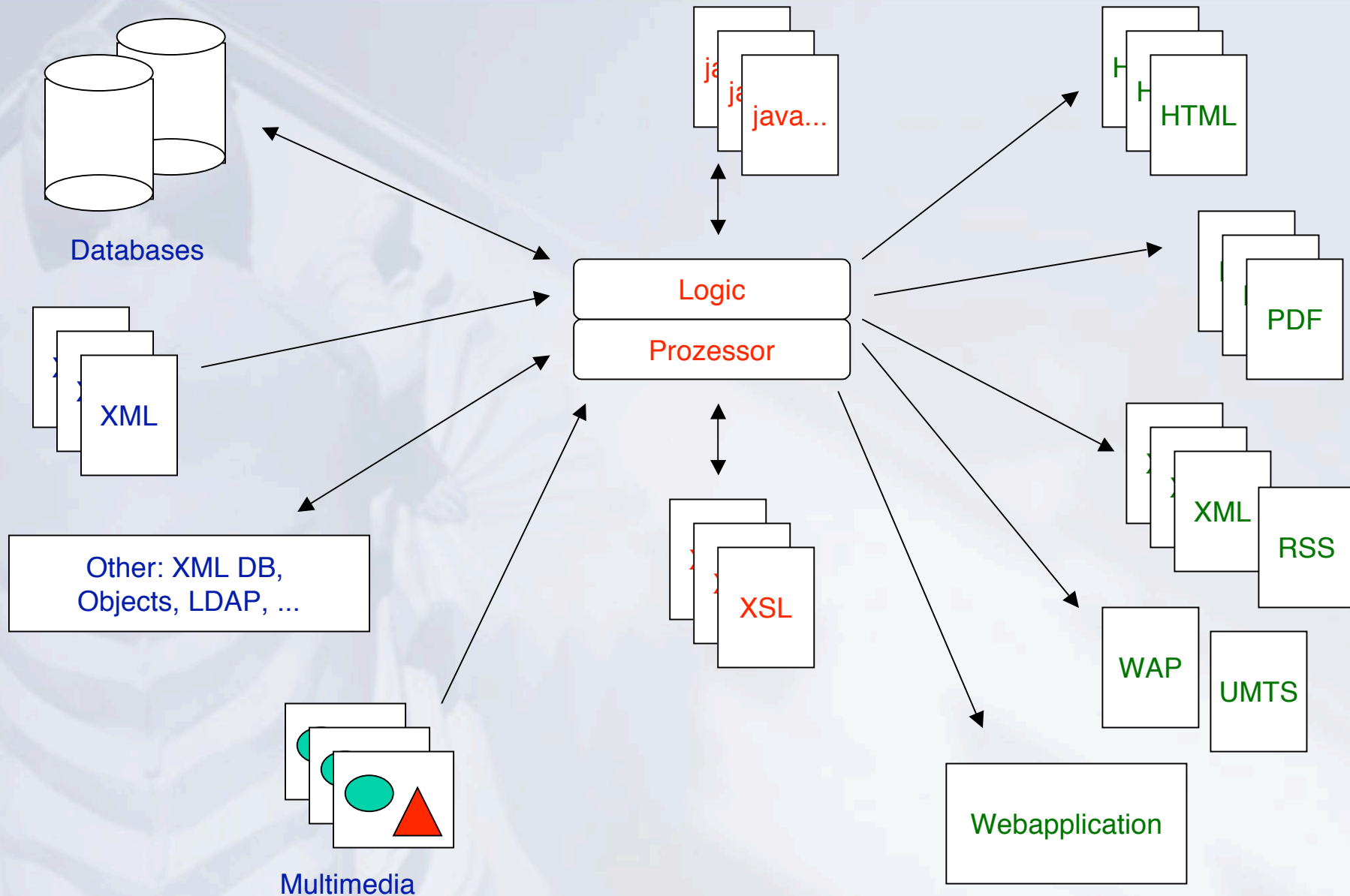
2. Presented Frameworks & Tools

3. Hands-On

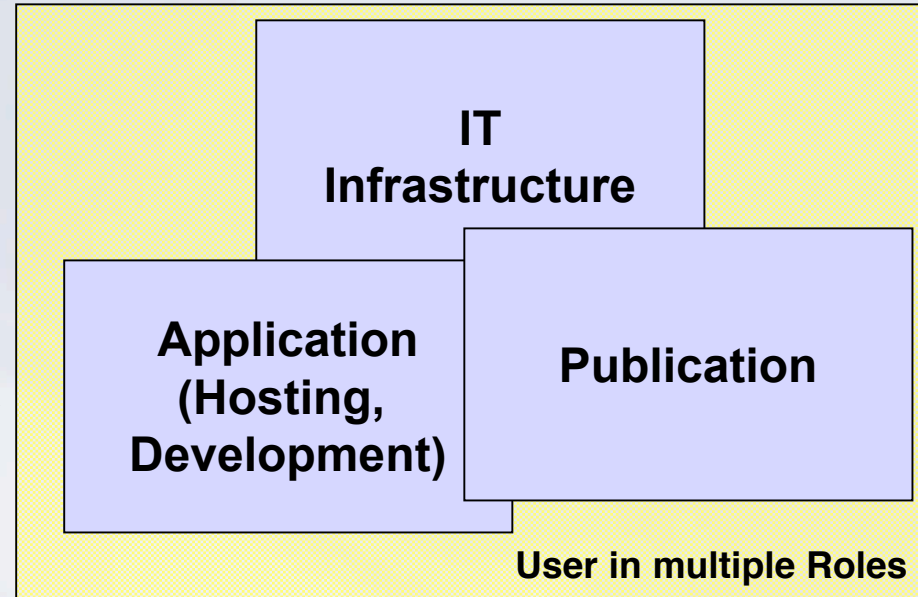
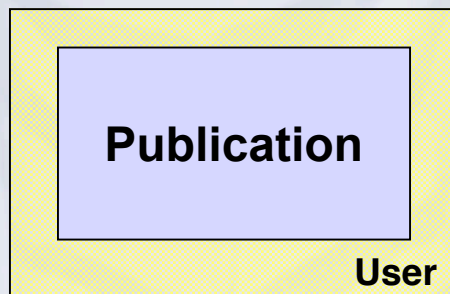
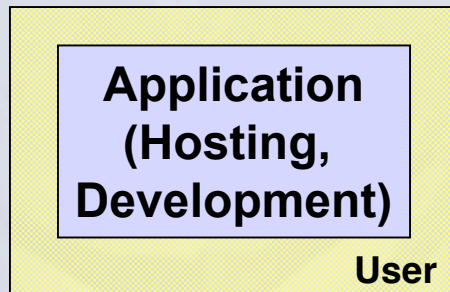
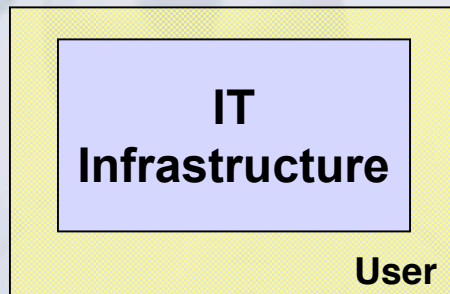
- Transformations Using Xalan (e.g., in JEdit)
- Batch Transformations using Apache Ant
- Online XML Processing using Apache Cocoon

1. Introduction

- **Cross-Publishing?**
- **Multi-Channel Publishing?**
- **System Integration**
- **Open Source Tools**
 - Importance
 - Capabilities



System Integration



- From „simple“ Publishing to Web Applications
- Intranet („thin client“) + Internet
- System Integration required
- Common Standards like Web-Services
- Content Management <--> Access to Enterprise Information Systems
- Publishing Integrated to Business Workflows

Data Sources (Examples)

- *XML File(s)*
- XML Streams (e.g., Network)
- Webservices
- *Relational Databases*
- XML Databases
- LDAP
- CVS, Subversion, ...
- Objects (O/R Mapping)
- SAP
- Multimedia (*JPEG*, GIF, PNG, SVG)

The *Marked Sources* are used in the Tutorial.

Target Formats (Examples)

- ***XML (e.g., RSS)***
- ***HTML***
- ***xHTML***
- **WML**
- ***JPEG*, PNG**
- **SVG**
- ***PDF*, RTF, Office**
- **(Relational) Databases...**

The *Marked Sources* are used in the Tutorial.

Open Source/Protocol

- **System and Vendor independent Protocols and Formats**
- **Standardisation (World Wide Web Consortium)**
- **Easy Data Exchange and System Integration**
- **No License Fees**
- **High Quality, Availability and Flexibility**
 - Different Languages like Java, Python, Perl, .net
 - Different Systems like Linux, BSD, OS X, Solaris, Windows, ...

2. Framework and Tools

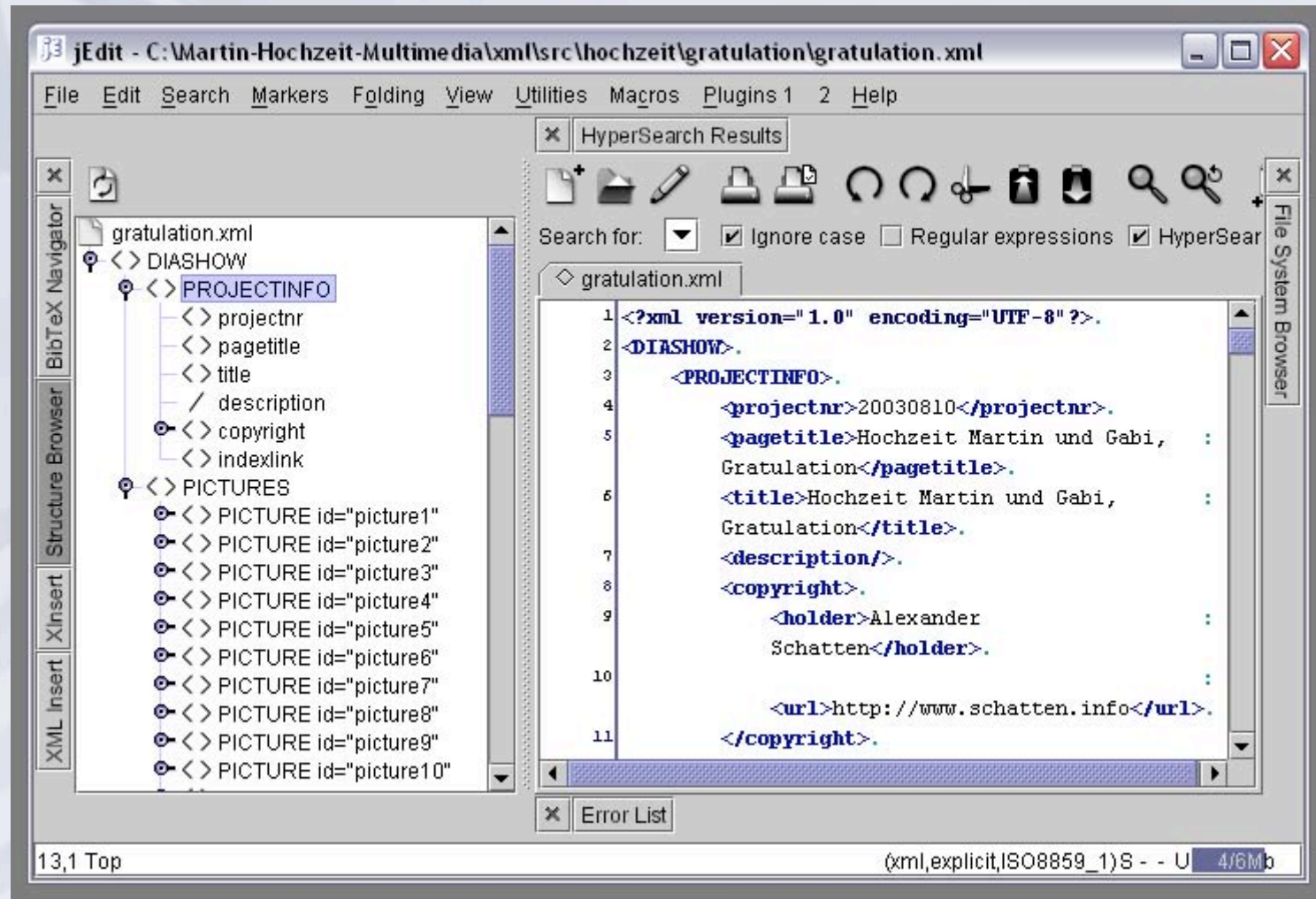
- **JEdit Editor, Eclipse**
- **Mozilla Browser**
- **Apache Xalan (XSLT Processor)**
- **Apache FOP (Formatting Objects Processor)**
- **Apache Ant (Batch Processing Tool)**
- **Apache Cocoon (XML Framework)**

JEdit



- **Java based**
 - Available for Mac, Linux, Windows, Solaris...
- **Powerful basic Edit Functions**
- **Flexible Plugin Concept with various Plugins**
 - Programming
 - XML
 - Project Organisation like CVS, ftp
 - ...
- **Supports many Formats like C++, Java, Perl, Python, XML, SQL, ...**

JEdit



Eclipse

- **Integrated Development Environment**
- **Java, C++ IDE**
- **Essentially: Plugin Framework**
- **Plugins**
 - XML
 - UML
 - CVS, Subversion
 - ...



Mozilla Browser



Apache Xalan

- **XSLT (Transformations) Processor**
- **Usage:**
 - Command Line
 - Java, C++ API
 - Servlet
- **W3C Standard Compatible**
- **Proprietary Extension Mechanism**
- **XSLTC (pre-compiled XSL)**

Apache FOP



- **Formatting Objects Processor**
- **Rather early Stage**
- **Still not fully Standard Compatible**
- **Bugs in some Areas like Tables**
- **Usage:**
 - Command Line
 - Java, C++ API
 - Servlet

Apache Ant



- **Platform Independent Batch Processing Tool (comparable to Unix make)**
- **Main Usage: Java Application Building**
 - Preparation
 - Compilation
 - Distribution
- **Many Tasks like Copy, Move, Zip, FTP, ...**
- **Usage for Batch XSLT Processing possible!**

Apache Cocoon



- **XML / XSL Publishing Framework**
- **Integration of many Apache Projects like: Xalan, Batik (SVG), FOP, Avalon, Xerces, ...**
- **Usage as Servlet for**
 - Multichannel Publishing
 - Application Server
 - System Integration
- **Usage as Command Line Tool**
- **Logic can be Implemented in**
 - XSLT, Java, XSP, JavaScript (Flow), SQL

Get the Tools: URLs

- **Used Tools**

- <http://www.jedit.org>
- <http://www.eclipse.org>
- <http://www.mozilla.org>
- <http://java.sun.com> (Java Virtual Machine)
- <http://ant.apache.org>
- <http://jakarta.apache.org> (Tomcat and other Proj.)
- <http://xml.apache.org> (Xerces, Xalan, FOP)
- <http://cocoon.apache.org>

- **Standards**

- <http://www.w3.org> (World Wide Web Consortium)

3. „Hands-On“

- **Transformations Using Xalan (in JEdit)**
- **Batch Transformations using Apache Ant**
- **Online XML Processing using Apache Cocoon**

Transformation Using Xalan (XML → XML)

```
<PERIODIC_TABLE>

  <ATOM>
    <NAME>Actinium</NAME>
    <ATOMIC_WEIGHT>227</ATOMIC_WEIGHT>
    <ATOMIC_NUMBER>89</ATOMIC_NUMBER>
    ...
  </ATOM>

  <ATOM>
    ...
    ...
  </ATOM>

  ...

</PERIODIC_TABLE>
```

```
<periodensystem>

  <atom
    siedepunkt=""
    atomgewicht="262"
    name="Bohrium"
  />

  <atom
    siedepunkt=""
    atomgewicht="247"
    name="Berkelium"
  />

  ...

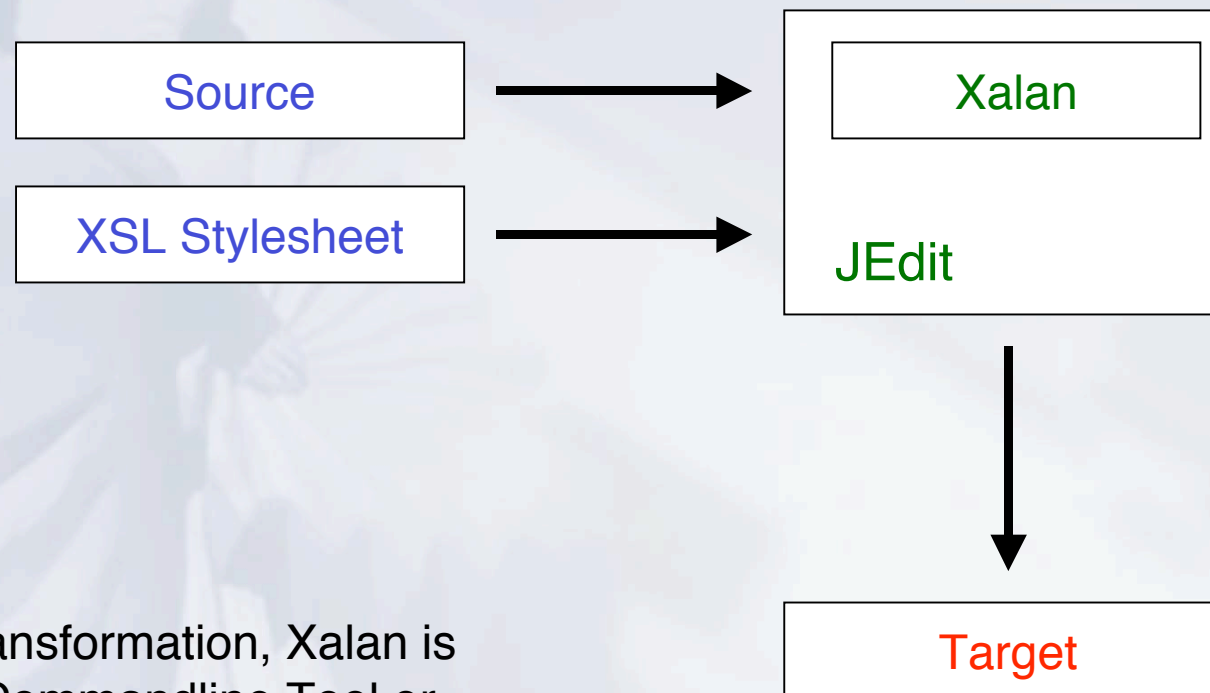
</periodensystem>
```

Source



Target

Transformation using Xalan

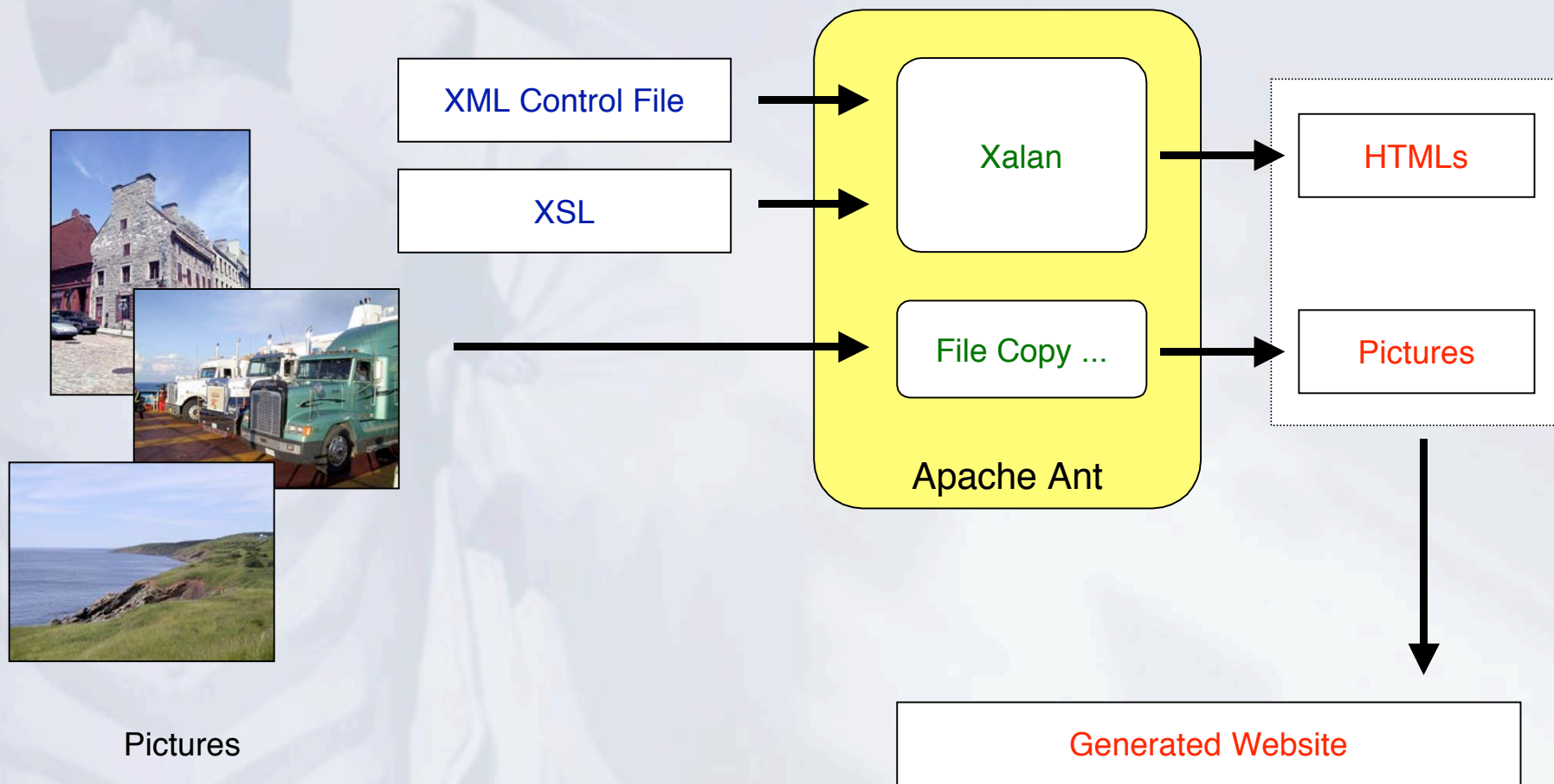


In this Transformation, Xalan is used as Commandline Tool or Embedded in an XML Editor

Batch Processing using Apache Ant

- **Example: Generate a Dia-Show:**
 - Source Images
 - Control (XML) File, that contains order, description of images ...
- **Dias-Show Website:**
 - Batch (offline) Generation
 - Automatic Navigation
 - Index Page
 - Variable Design

Batch Processing using Apache Ant



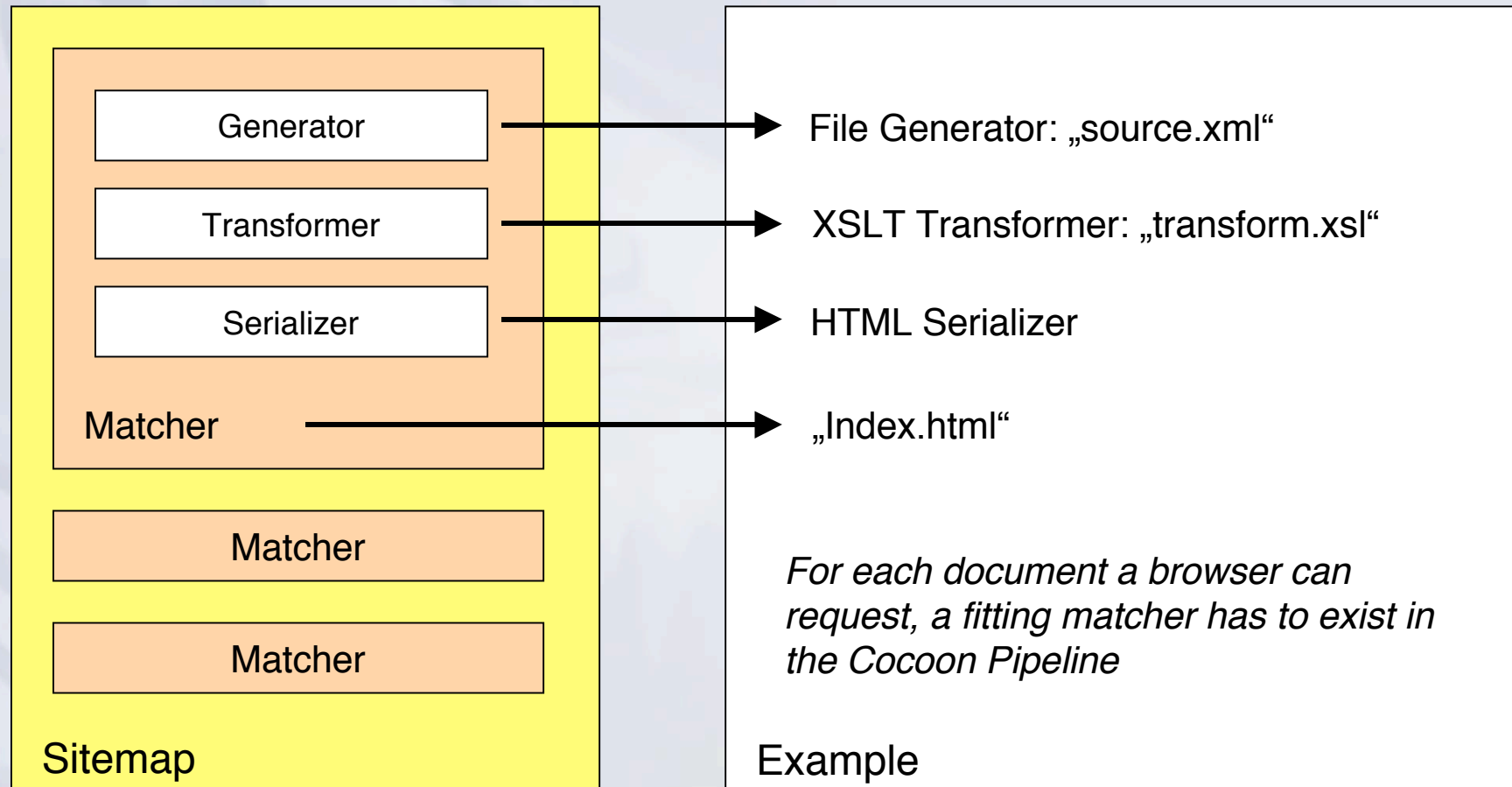
Online Publishing using Cocoon

- **Render Information from XML File**
 - In HTML
 - In PDF
- **Publish Information from Relational Database**
- **User Interaction using Forms**
- **Pipeline/Matcher Concepts**
- **Brief Overview over advanced Concepts like**
 - Session Handling
 - Authentication
 - Caching
 - Flow

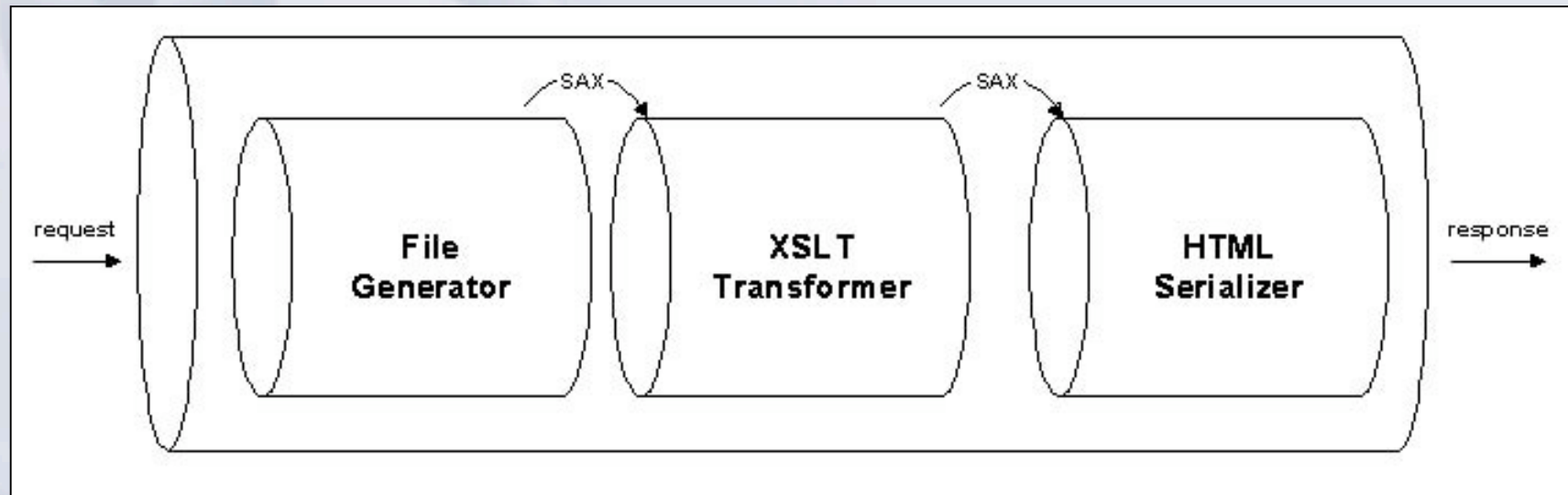
Cocoon Basic Ideas

- **A Cocoon Application is based on a „Sitemap“**
- **Each Document to be Published is Declared in the Sitemap**
- **Various Components can be used in the Sitemap like**
 - Generators
 - Transformers (XSLT, SQL, Castor, ...)
 - Actions
 - XSP (JSP like) Logic
 - Serializers
 - Database Components

Cocoon Sitemap: „Virtual Documents“



Cocoon Pipeline: Simple Example



1. This Example shows a File Generator that reads the XML Document
2. The XSLT Transformer generates XHTML
3. The HTML serializer sends the HTML document to the Browser

Cocoon Example

- **Among others, the following Cocoon mechanisms are illustrated in the „Hands-On“ Examples:**
 - „Simple“ XML → HTML Transformations
 - Generation of HTML and PDF
 - Matchers that match to a series of documents
 - Access to Relational Databases
 - Simple Form Handling
 - (Simple Sitemap based Authentication)

Moreover...

- **Form Framework**
- **Internationalisation**
- **Portal Framework (Java Portlets)**
- **Session Handling**
- **Authentication**
- **...**

Contact

For detailed Information or Questions please contact

Dr. Alexander Schatten
Institute for Software Technology
and Interactive Systems
Vienna University of Technology

alexander@schatten.info
<http://www.schatten.info>
<http://www.ifs.tuwien.ac.at>

**Additionally, the „Hands-On“ Examples
from this Tutorial can be found on my Website**