Welcome to

3. SOAP, WSDL, XML data

KEA System Integration F2020 10 ECTS

Henrik Lund Kramshøj hlk@zencurity.com @kramse

Slides are available as PDF, kramse@Github

3-SOAP-WSDL-XML-system-integration.tex in the repo security-courses

Plan for today

- Data overview
- XML data, JSON
- Data Transformation
- WebServices
- SOAP, WSDL
- Investigate examples

Exercises

-
-

Reading Summary

Camel chapter 3: Transforming data with Camel

Contains a lot of different formats

Data overview XML data, JSON

XML data

> Extensible Markup Language (XML) is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable. The World Wide Web Consortium's XML 1.0 Specification[2] of 1998[3] and several other related specifications[4]—all of them free open standards—define XML.[5]
> The design goals of XML emphasize simplicity, generality, and usability across the Internet.[6] It is a textual data format with strong support via Unicode for different human languages. Although the design of XML focuses on documents, the language is widely used for the representation of arbitrary data structures[7] such as those used in web services.

Source: `https://en.wikipedia.org/wiki/XML`

- We have seen XML used for configuration in Apache Tomcat and Camel
- Perfect for computers, less for humans

XML data example - Nmap output

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE nmaprun>
<?xml-stylesheet href="file:///usr/bin/../share/nmap/nmap.xsl" type="text/xsl"?>
<!-- Nmap 7.70 scan initiated Sat Feb 22 23:35:53 2020 as: nmap -oA router -sP 10.0.42.1 -->
<nmaprun scanner="nmap" args="nmap -oA router -sP 10.0.42.1" start="1582410953"
 startstr="Sat Feb 22 23:35:53 2020" version="7.70" xmloutputversion="1.04">
<verbose level="0"/>
<debugging level="0"/>
<host><status state="up" reason="echo-reply" reason_ttl="62"/>
<address addr="10.0.42.1" addrtype="ipv4"/>
<hostnames>
</hostnames>
<times srtt="2235" rttvar="5000" to="100000"/>
</host>
<runstats><finished time="1582410953" timestr="Sat Feb 22 23:35:53 2020" elapsed="0.32"
 summary="Nmap done at Sat Feb 22 23:35:53 2020; 1 IP address (1 host up)
 scanned in 0.32 seconds" exit="success"/><hosts up="1" down="0" total="1"/>
</runstats>
</nmaprun>
```

XML data - documents

> Hundreds of document formats using XML syntax have been developed,[8] including RSS, Atom, SOAP, SVG, and XHTML. XML-based formats have become the default for many office-productivity tools, including Microsoft Office (Office Open XML), OpenOffice.org and LibreOffice (OpenDocument), and Apple's iWork[citation needed]. XML has also provided the base language for communication protocols such as XMPP. Applications for the Microsoft .NET Framework use XML files for configuration, and property lists are an implementation of configuration storage built on XML.[9]

Source: `https://en.wikipedia.org/wiki/XML`

- Document formats using XML may still be proprietary!
- Documents using XML can be validated, are they well-formed according to the Document Type Definition (DTD)

XML data - standards

> Many industry data standards, such as Health Level 7, OpenTravel Alliance, FpML, MISMO, and National Information Exchange Model are based on XML and the rich features of the XML schema specification. Many of these standards are quite complex and it is not uncommon for a specification to comprise several thousand pages.[citation needed] In publishing, Darwin Information Typing Architecture is an XML industry data standard. XML is used extensively to underpin various publishing formats.

Source: `https://en.wikipedia.org/wiki/XML`

-
-

XML data for Service-oriented architecture (SOA)

> XML is widely used in a Service-oriented architecture (SOA). Disparate systems communicate with each other by exchanging XML messages. The message exchange format is standardised as an XML schema (XSD). This is also referred to as the canonical schema. XML has come into common use for the interchange of data over the Internet. IETF RFC:3023, now superseded by RFC:7303, gave rules for the construction of Internet Media Types for use when sending XML. It also defines the media types application/xml and text/xml, which say only that the data is in XML, and nothing about its semantics.

Source: `https://en.wikipedia.org/wiki/XML`

- We will talk more about SOA later

Transforming XML using XSLT

> XSLT (Extensible Stylesheet Language Transformations) is a language for transforming XML documents into other XML documents,[1] or other formats such as HTML for web pages, plain text or XSL Formatting Objects, which may subsequently be converted to other formats, such as PDF, PostScript and PNG.[2] XSLT 1.0 is widely supported in modern web browsers.[3]

Source: `https://en.wikipedia.org/wiki/XSLT`

-
-

xsltproc example using Nmap

```
$ su -
# apt install nmap xsltproc
# nmap -sP -oA /tmp/router 91.102.91.18
# exit
$ xsltproc /tmp/router.xml > /tmp/router.html
$ firefox /tmp/router.html
```

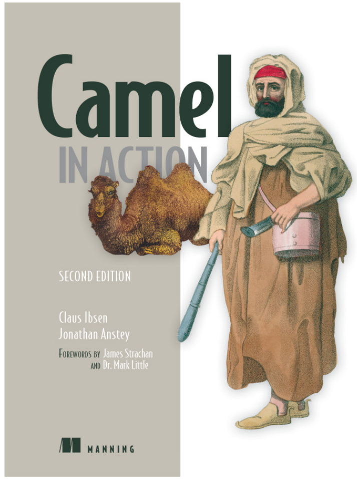
- We can use the command line tool `xlstproc` for transforming documents
- `apt install xsltproc`
- Its part of the package Libxslt `https://en.wikipedia.org/wiki/Libxslt`

Data oveview JSON

Source:

-
-

Review Still got the Debian VM running



Chapter 1: file-copy example

```
hlk@debian-lab:~/projects/system-integration/camelinaction2/chapter1/file-copy$ find data/
data/
data/outbox
data/outbox/message1.xml
data/inbox
data/inbox/message1.xml
```

- We want to run the command for Maven to download tools, and do stuff
- `mvn compile exec:java`
- This might take some time!
- Note: this is a two step process, so split into `mvn compile` and `exec:java` if you have trouble running

Success Compile

```
hlk@debian-lab:~/projects/system-integration/camelinaction2/chapter1/file-copy$ mvn compile
[INFO] Scanning for projects...
[INFO]
[INFO] ----------------< com.camelinaction:chapter1-file-copy >----------------
[INFO] Building Camel in Action 2 :: Chapter 1 :: File Copy Example 2.0.0
[INFO] --------------------------------[ jar ]---------------------------------
[INFO]
[INFO] --- maven-resources-plugin:2.4.3:resources (default-resources) @ chapter1-file-copy ---
[INFO] Using 'UTF-8' encoding to copy filtered resources.
[INFO] Copying 1 resource
[INFO]
[INFO] --- maven-compiler-plugin:3.6.1:compile (default-compile) @ chapter1-file-copy ---
[INFO] Nothing to compile - all classes are up to date
[INFO] ------------------------------------------------------------------------
[INFO] BUILD SUCCESS
[INFO] ------------------------------------------------------------------------
[INFO] Total time: 1.270 s
[INFO] Finished at: 2020-02-17T07:08:02+01:00
[INFO] ------------------------------------------------------------------------
```

Success Execute Java - shortened for slide

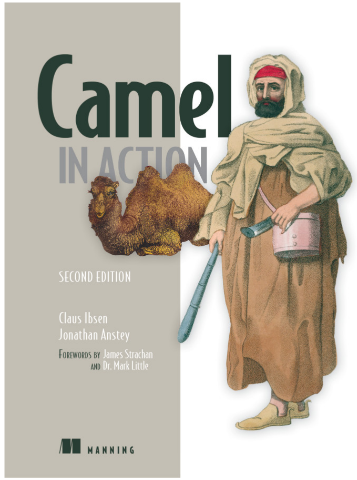
```
hlk@debian-lab:~/projects/system-integration/camelinaction2/chapter1/file-copy$ mvn exec:java
[INFO] Scanning for projects...
[INFO]
[INFO] ----------------< com.camelinaction:chapter1-file-copy >----------------
[INFO] Building Camel in Action 2 :: Chapter 1 :: File Copy Example 2.0.0
[INFO] --------------------------------[ jar ]---------------------------------
[INFO]
[INFO] --- exec-maven-plugin:1.2.1:java (default-cli) @ chapter1-file-copy ---
[ion.FileCopierWithCamel.main()] DefaultCamelContext  INFO  Apache Camel 2.24.3 (CamelContext: camel-1) is starting
[ion.FileCopierWithCamel.main()] FileEndpoint         INFO  Using default memory based idempotent repository with cache max size: 10
[ion.FileCopierWithCamel.main()] DefaultCamelContext  INFO  Route: route1 started and consuming from: file://data/inbox?noop=true
[ion.FileCopierWithCamel.main()] DefaultCamelContext  INFO  Total 1 routes, of which 1 are started
[INFO] ------------------------------------------------------------------------
[INFO] BUILD SUCCESS
[INFO] ------------------------------------------------------------------------
[INFO] Total time: 11.908 s
[INFO] Finished at: 2020-02-17T07:11:18+01:00
[INFO] ------------------------------------------------------------------------
```

Success Execute Java - new files

```
$ find data/
data/
data/outbox
data/outbox/message1.xml
data/outbox/message2.txt
data/inbox
data/inbox/message1.xml
data/inbox/message2.txt
```

- Try adding a new file using editor, and re-run
  `echo "some data" > data/inbox/message2.txt`

Chapter 3: Transforming data with Camel



Camel

IN ACTION

SECOND EDITION

Claus Ibsen
Jonathan Anstey

Forewords by James Strachan
and Dr. Mark Little

MANNING

This chapter covers

- Transforming data by using EIPs and Java
- Transforming XML data
- Transforming by using well-known data formats

- Writing your own data formats for transformations
- Understanding the Camel type-converter mechanism

3.1 Data transformation overview

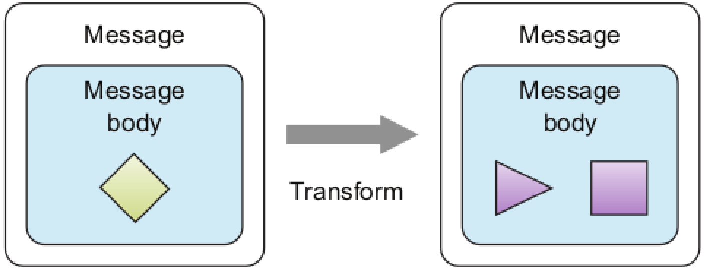


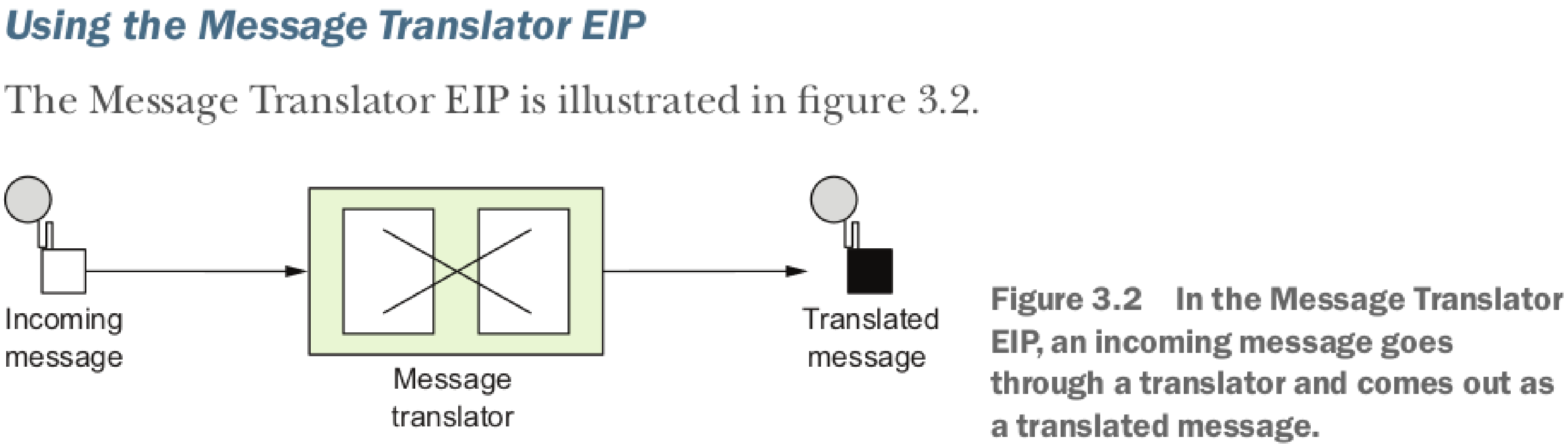
Figure 3.1 Camel offers many features for transforming data from one form to another.

- Data format transformation – The data format of the message body is transformed from one form to another. For example, a CSV record is formatted as XML.
- Data type transformation – The data type of the message body is transformed from one type to another. For example, java.lang.String is transformed into javax. jms.TextMessage .

Six ways data transformation typically takes place in Camel

- Data transformation using EIPs and Java
- Data transformation using components
- Data transformation using data formats
- Data transformation using templates
- Data type transformation using Camel's type-converter mechanism
- Message transformation in component adapters

3.2 Message Translator EIP



Using the Message Translator EIP

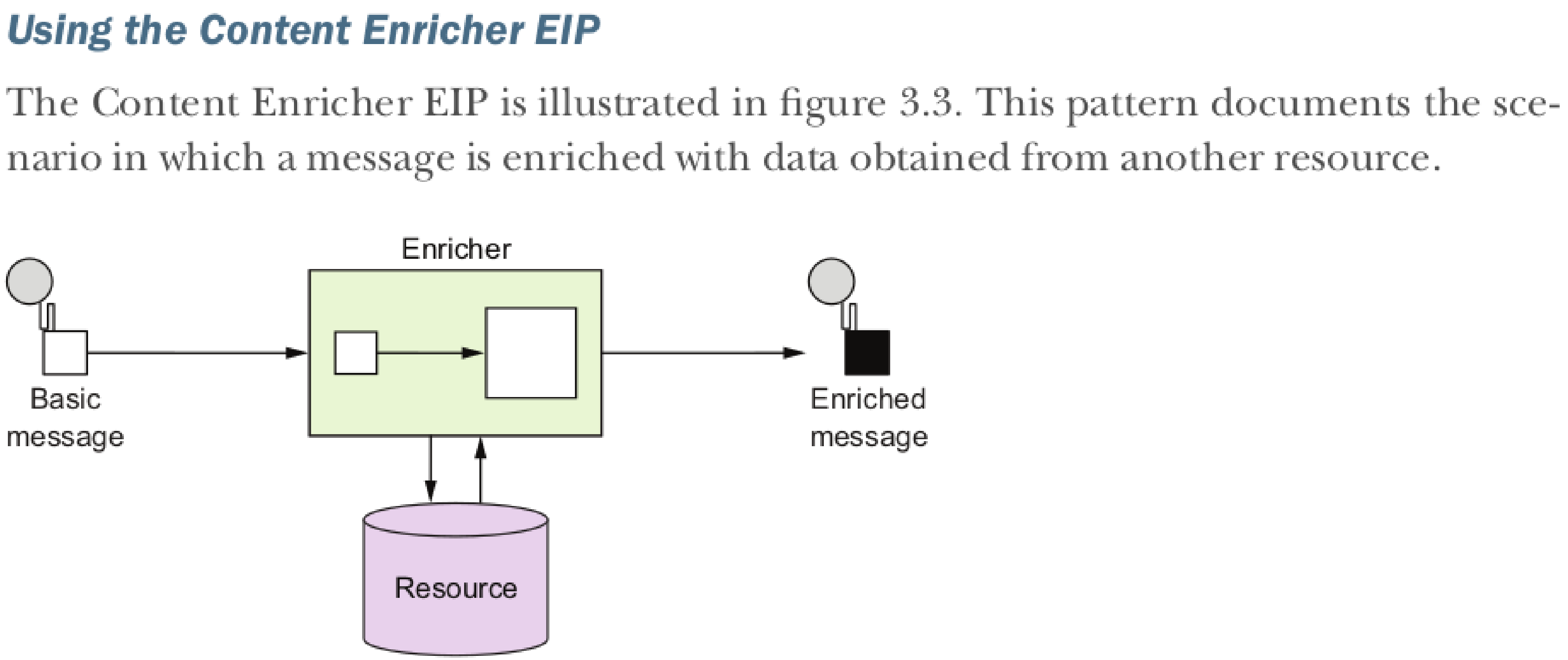
The Message Translator EIP is illustrated in figure 3.2.

Incoming message
Message translator
Translated message

Figure 3.2 In the Message Translator EIP, an incoming message goes through a translator and comes out as a translated message.

- 3.2 Transforming data by using EIPs and Java
- Data mapping, the process of mapping between two distinct data models, is a key fac- tor in data integration.
- Picture shown is using the Message Translator EIP

3.2 Content Enricher EIP



Using the Content Enricher EIP

The Content Enricher EIP is illustrated in figure 3.3. This pattern documents the scenario in which a message is enriched with data obtained from another resource.

Enricher
Basic message
Enriched message
Resource

- Using the Content Enricher EIP

3.3 Transforming XML

- Transforming XML with XSLT
- Transforming XML with object marshaling

3.4 Transforming with data formats

- Data formats provided with Camel
- Using Camel's CSV data format
- Using Camel's Bindy data format
- Using Camel's JSON data format
- Configuring Camel data formats

3.5 Transforming with templates

- Using Apache Velocity

3.6: Understanding Camel type converters

- How the Camel type-converter mechanism works
- Using Camel type converters
- Writing your own type converter

3.7 Summary and best practices

Web services SOAP, WSDL,

Web Services

- Today a generic name for services using the internet
- Web servers such as Apache HTTPD, Nginx etc. provide a service to thew internet allowing access using HTTP
- Source for some parts on this slide, `https://en.wikipedia.org/wiki/Web_service`

SOAP - Simple Object Access Protocol

> SOAP (abbreviation for Simple Object Access Protocol) is a messaging protocol specification for exchanging structured information in the implementation of web services in computer networks. Its purpose is to provide extensibility, neutrality, verbosity and independence. It uses XML Information Set for its message format, and relies on application layer protocols, most often Hypertext Transfer Protocol (HTTP), although some legacy systems communicate over Simple Mail Transfer Protocol (SMTP), for message negotiation and transmission.

Source: `https://en.wikipedia.org/wiki/SOAP`

WSDL - Web Services Description Language

The Web Services Description Language (WSDL / w z dəl/) is an XML-based interface description language that is used for describing the functionality offered by a web service. The acronym is also used for any specific WSDL description of a web service (also referred to as a WSDL file), which provides a machine-readable description of how the service can be called, what parameters it expects, and what data structures it returns. Therefore, its purpose is roughly similar to that of a type signature in a programming language.

The current version of WSDL is WSDL 2.0. The meaning of the acronym has changed from version 1.1 where the "D"stood for "Definition".

Source: `https://en.wikipedia.org/wiki/Web_Services_Description_Language`

For Next Time

Think about the subjects from this time, write down questions

Check the plan for chapters to read in the books

Visit web sites and download papers if needed

Retry the exercises to get more confident using the tools