XML: Summary and conclusions

§ XML standards

§ Software development and application interfaces

§ Data Base Management case: Oracle

§ Case: Microsoft .NET

§ Web Services

§ Semantic Web

Conclusions: XML standards 1

§ XML = Extended Markup Language

   § a meta language for the creation of languages to define document structures

§ XSL = XML Stylesheet Language

   - XSLT = XSL Transformations
     - a transformation language to transform XML structures to other XML structures, HTML or text
   - FO = Formatting Objects (or XSL-FO)
     - a style language for XML document layout on paper or in electronic format

Conclusions: XML standards 2

§ Xpath = XML Path Language
  § a language for navigation of XML documents and locating elements
  § XPointer
  § Xlink
§ Namespaces in XML
  § unambiguous naming of elements and attributes
§ XML Schema
§ XHTML
§ XForms 1.0 recommendation 14.10.03
§ XQuery 1.0 recommendation 2007

XML application areas

§ Multimedia
  - § Voice XML
  - § SVG 1.1
  - § SMIL 2.0
  - § X3D and x-smiles (Helsinki Univ. of Technology)

§ electronic commerce, EDI (Electronic Data Interchange)
  - § ebXML: electronic business XML, UN/CEFACT and OASIS
    - • main elements: ebXML storage, CPP – Collaboration Protocol Profile, CPA – Collaboration Partner agreement
  - § BizTalk
  - § RosettaNet & PapiNet

§ HR-XML Human Resources management system for data exchange

XML application areas

§ Publishing, multichannel publishing
  - § documents, DocBook
  - § metadata
§ Documentation:
  - § technical documentation: manuals, term banks, spare part catalogs, language versions
  - § for example Dublin Core
§ Reusability:
  - § same information can be deliverd through different media: multichannel publishing
  - § WWW, PDA, mobile, DVD, print

Software development approaches
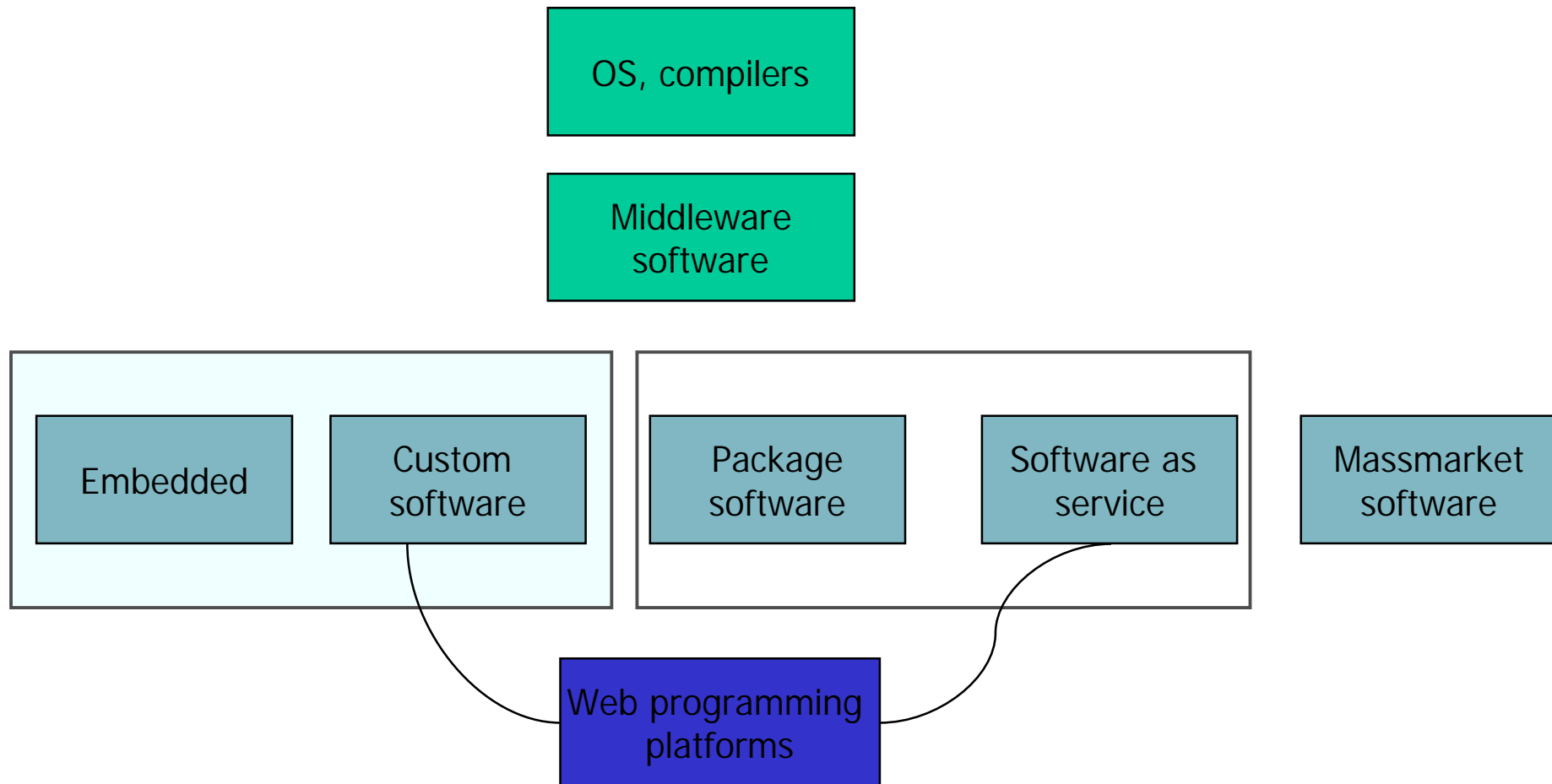
Service-Oriented Architecture SOA

§ Service-Oriented Architecture (SOA) facilitates the development of modular business services that can be easily integrated and reused—creating a truly flexible, adaptable IT infrastructure.

§ CORBA, Web Services

§ Distributed computing, cloud computing

§ Modular programming

§ SOA and Business Architecture; a mechanism for defining business services

§ SAP Enterprise Services Architecture

§ Oracle, Accenture, etc.

Application design and UI languages

§ XUL: Firefox & Thunderbird

§ Silverlight for Designers Microsoft

  § "XAML browser application is a powerful declarative markup language that is the foundation for creating engaging graphics, animation, and media in Silverlight. XAML is similar to HTML, but it is more powerful and extensible."

  § for mobile devices as well

§ Macromedia MXML

§ XAMJ, Open Source, Java based, clientnet architecture

§ XForms (W3C)

XML application interfaces

- § SAX = Simple API for XML
  - § event-based programming interface for XML documents
- § DOM = Document Object Model
  - § XML documents can be manipulated in object oriented software using Document Object Model
  - § The basic idea in DOM is to read XML documents into trees and manipulate the tree
  - § DOM is a W3C (www.w3.org) recommendation
- § (DOM) interfaces:
  - § programming language and implementation
  - § independent interfaces available for several languages including Java, C++, Python, and IDL

DOM

§ DOM is also used in browsers to manipulate HTML content

§ DOM levels

   § DOM Level 1

      • basic manipulation interfaces: application programming interface (API) for XML (and HTML) documents using tree-structure

   § DOM Level 2

      • builds on DOM1, adds interfaces for document navigation, CSS rules and event based XML document manipulation

   § DOM Level 3

      • XML Schemas, Xpath

A pipeline for tranformations

Jaana Holvikivi
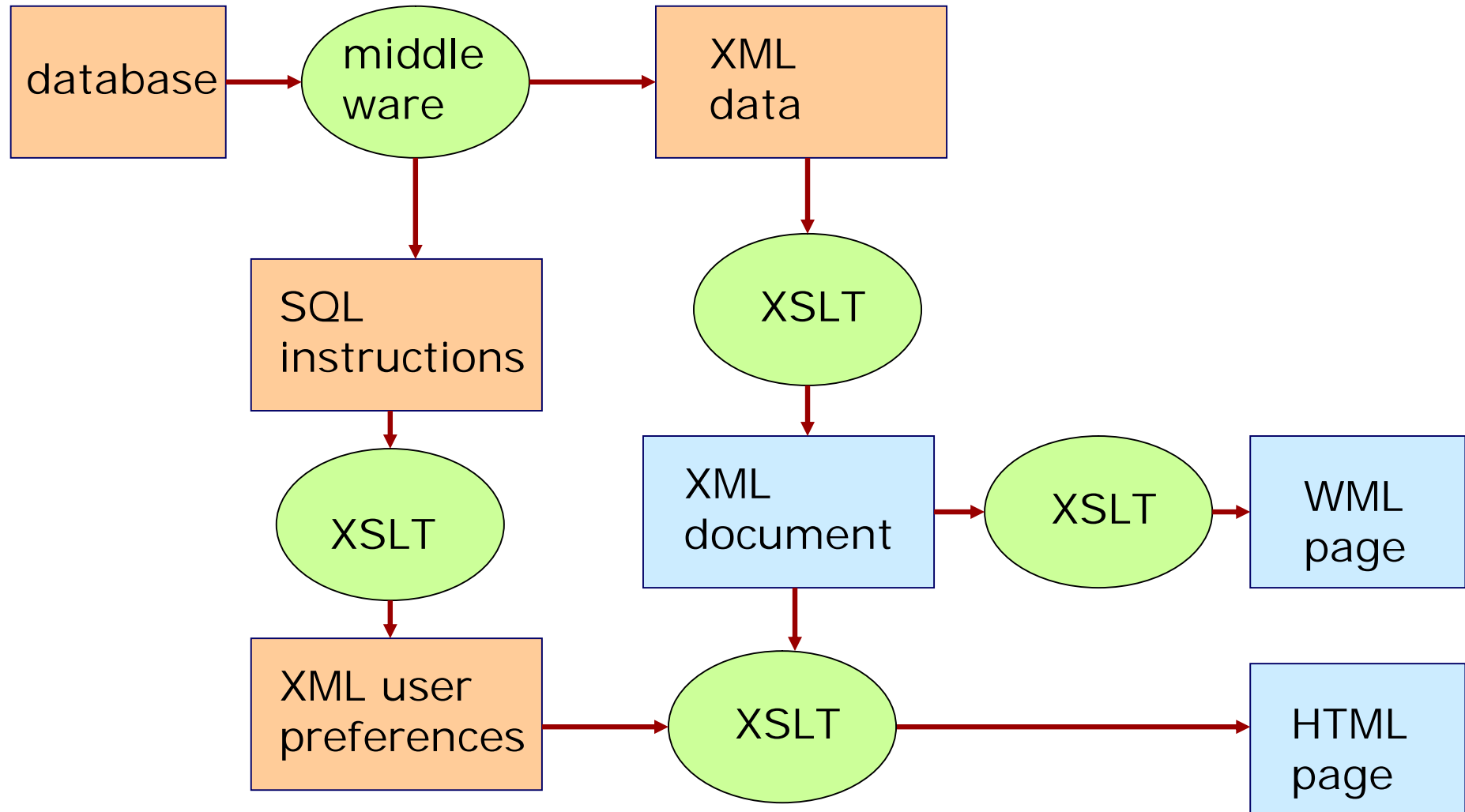
XML in data exchange and document storage

§ Data exchange and transfer: data-centered approach
  - § data transfer in XML format
  - § SOAP requests
  - § the order of data elements is not important
§ Storage: document-centered approach
  - § XML format
  - § i.e. SGML, XHTML, DocBook, news data bases
  - § the order of data elements is important
  - § designed to be read also by humans

Tools for different purposes

§ what is the main application of XML: storage or data transfer?

§ data transfer:

  § XML is generated from a data base

  § XML documents are entered into tables

  § off-the-shelf tools or design of a tailored application

§ XML functions as a document data base

  § are changes, searches or deletions needed?

  § versions?

Case: Oracle DBMS and XML

§ relational data base management system

§ Oracle 8i

  § basic XML generation and processing

  § XML developer's kit (XDK)

§ Oracle 9i R1

  § XML storage and searches integrated into data base

§ Oracle 9i R2

  § native XML DB

§ XDK

  § support for programming: Java, C, C++, PL/SQL

Case: Oracle DBMS and XML 2

§ XML DB: XML documents are processed in the data base
  - § generation
  - § validation
  - § XSLT transformations
  - § searches, updates, deletions
  - § not a separate server but a collection of XML technologies added to the Oracle data base
  - § inbuilt XML repository (directories)
  - § searches and interfaces: SQL/ Java, WebDAV, FTP, HTTP

Case : Microsoft .NET

Microsoft .NET has been released in late 2001 and covers the environment for application development with an emphasis for solutions over the internet.

§ in 2008 .NET Framework 3.5
§ .NET (dotnet) basic facts:
  § Visual Studio .NET
    • C++, C#, Visual Basic, based on CLR (Common Language Runtime)
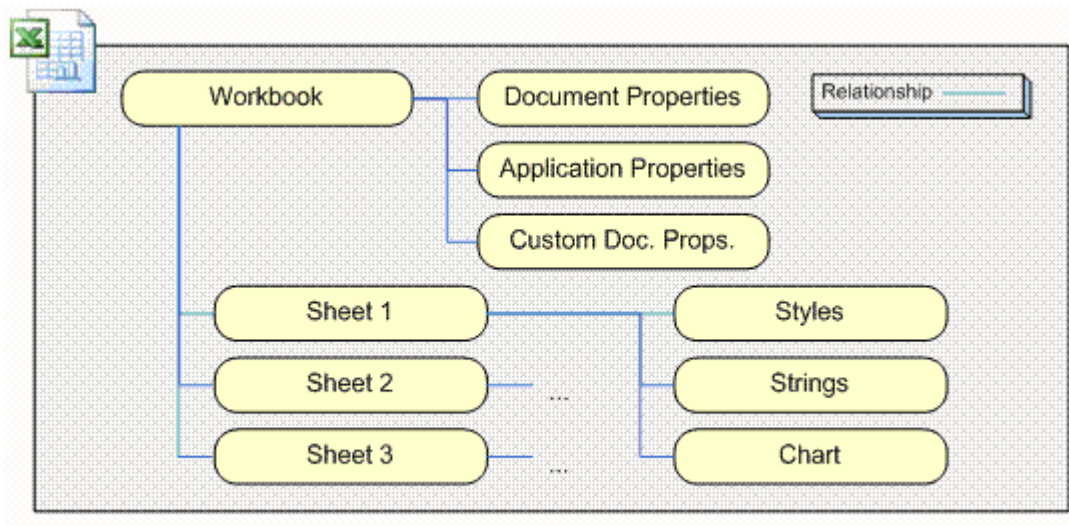 § Part of XP, 2003, Server 2008 & Vista operating systems
  § ADO data base interface implements XML
  § ASP.NET 2 (incl. Ajax)
  § Web Services based on XML
  § Most XML standards will be implemented (not Xlink?) except SAX, only DOM supported
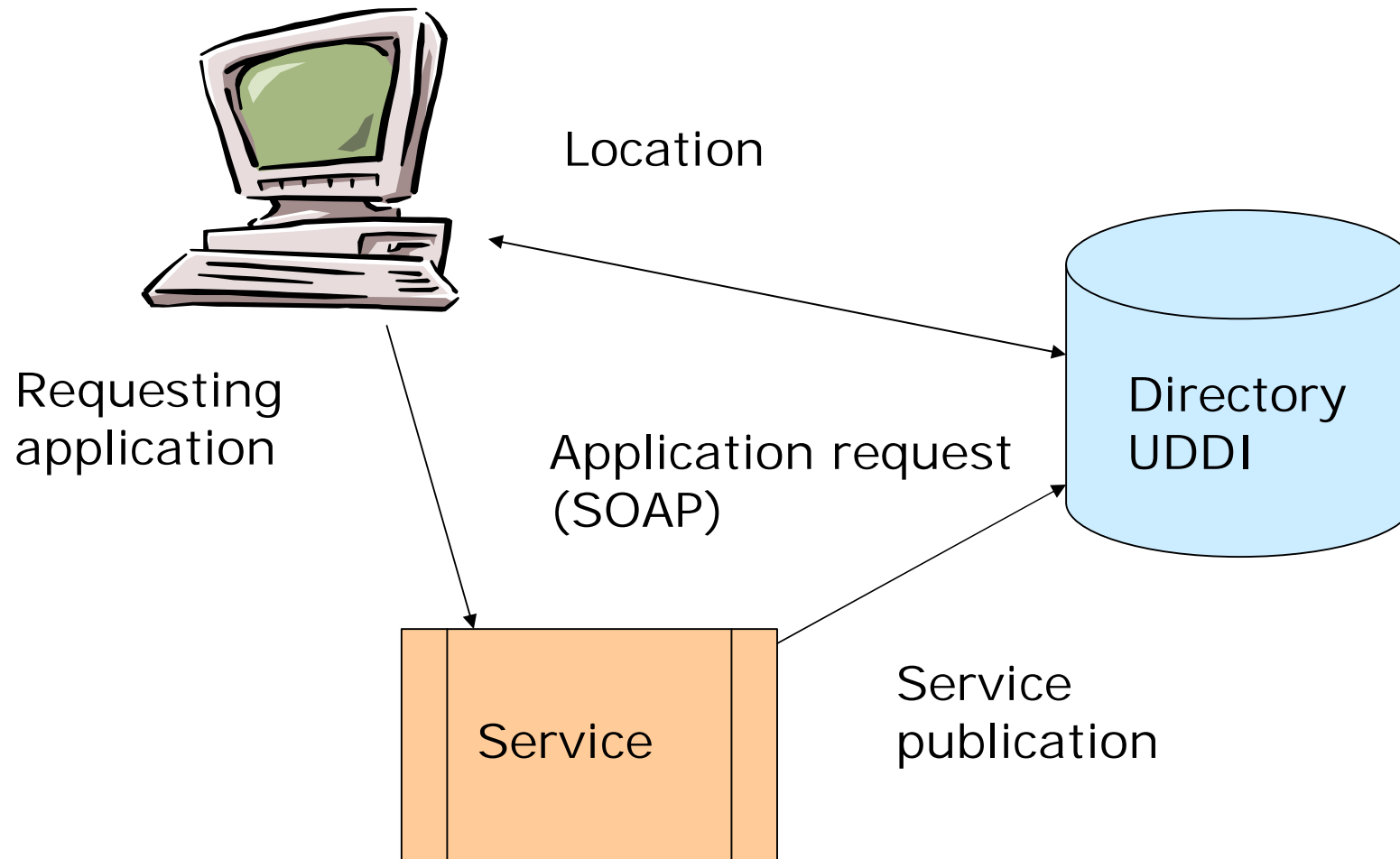
Office Open XML File Format

Microsoft architecture

§ WPF Windows Presentation Foundation
  - § in XML format, vector based graphics
  - § Graphics controls on desktop, animated user interfaces
  - § Silverlight is based on WPF but runs on a Web browser

§ WCF Windows Communication Foundation
  - § Unified communications that link TCP/IP, WebServices, HTTP, etc.

§ WWF Windows Workflow Foundation
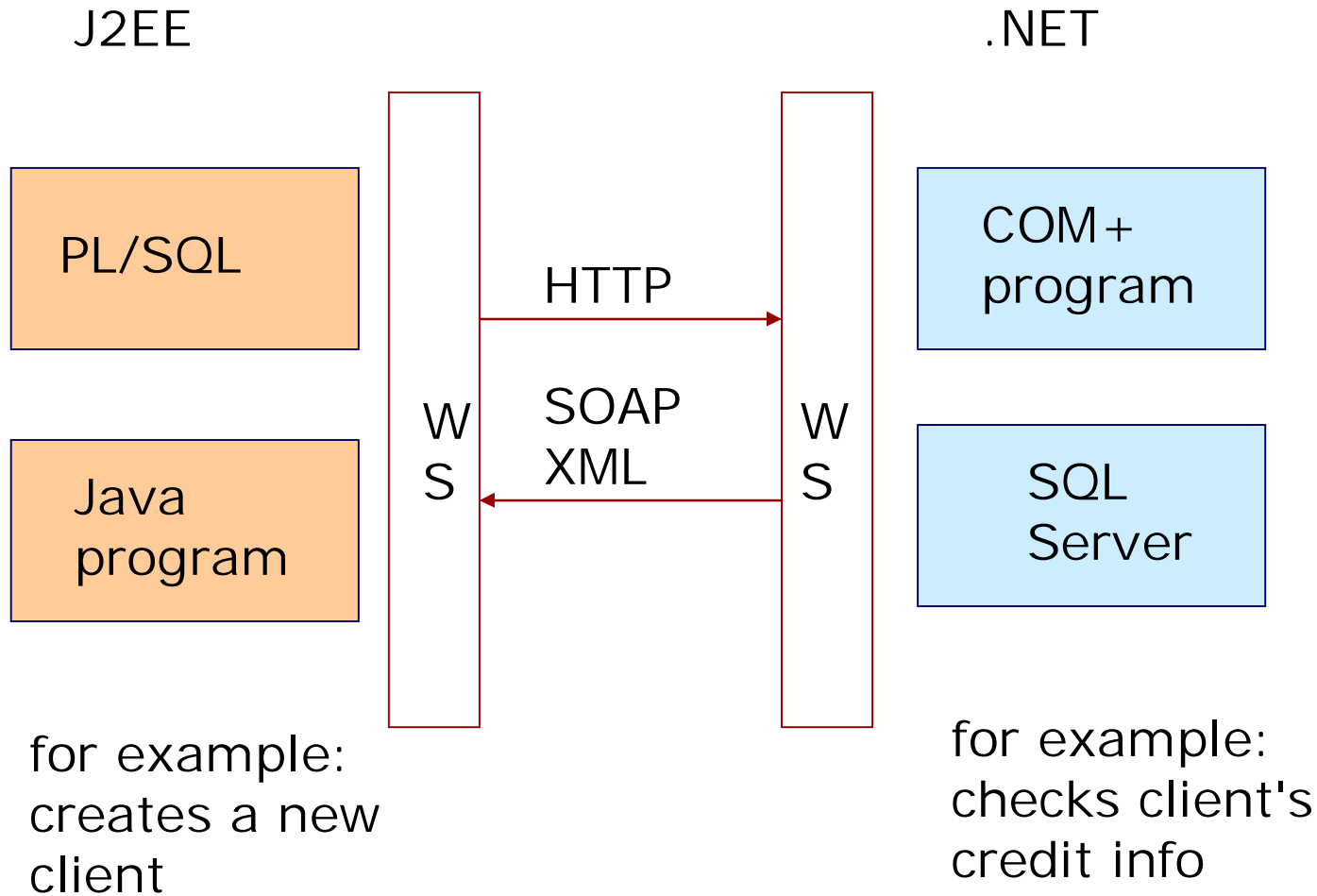 - § BizTalk Server
  - § Orchestrates workflows

Web Services

§ are functional components or programs, format XML

§ used over the internet (user server asks service component from another server and embeds it in the response file to the client)

§ key components

   § SOAP Simple Object Access Protocol

   § HTTP

   § Universal Description, Discovery, and Integration (UDDI)

   § WSDL Web Service Description Language

Web Services concept

Location

Requesting
application

Directory
UDDI

Application request
(SOAP)

Service

Service
publication

Web Services between platforms

J2EE

.NET

PL/SQL

COM+
program

HTTP

W
S

SOAP
XML

W
S

Java
program

SQL
Server

for example:
creates a new
client

for example:
checks client's
credit info

SOAP message

SOAP message

HTTP headers

SOAP envelope

SOAP header

SOAP body

XML encoded SOAP message name and data

SOAP example

```
<SOAP-ENV: Envelope>
 <SOAP-ENV:Body>
        <GetStockQuote
                xmlns:"urn:stock-quotes">
        <StockSymbol>ORCL</StockSymbol>
        </GetStockQuote>
 </SOAP-ENV:Body>
</SOAP-ENV: Envelope>
```
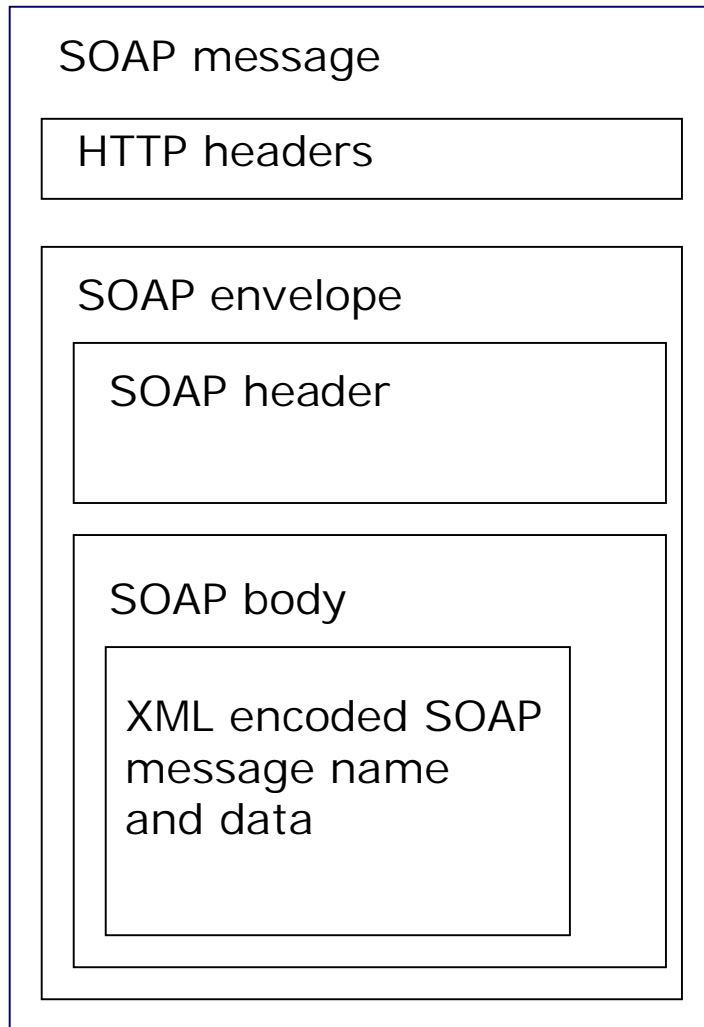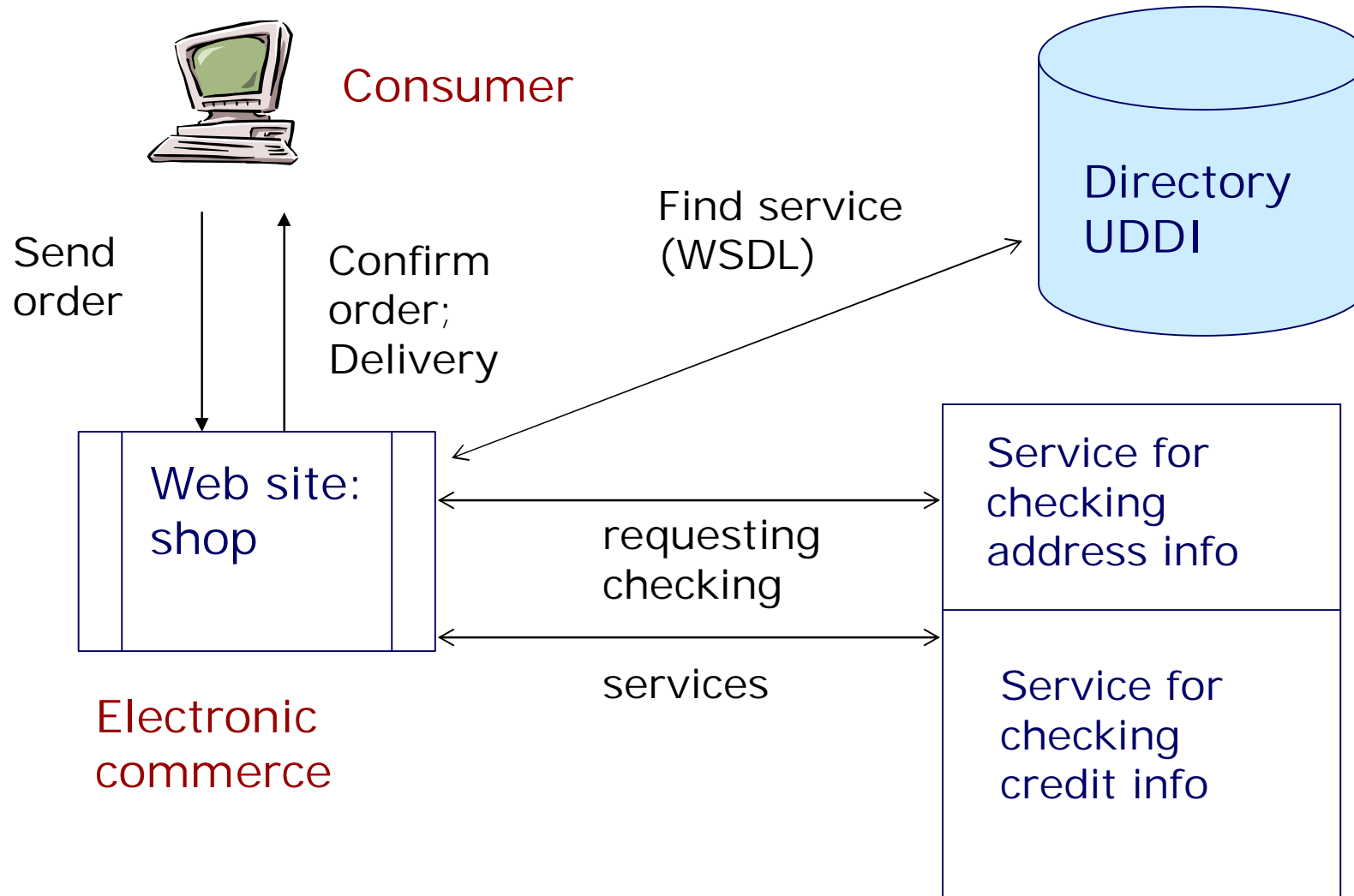
Request

```
<SOAP-ENV: Envelope>
 <SOAP-ENV:Body>
        <GetStockQuoteResponse
                xmlns:"urn:stock-quotes">
        <USD_Price>17.18</USD_Price>
        </GetStockQuoteResponse>
 </SOAP-ENV:Body>
</SOAP-ENV: Envelope>
```

Response

Web Services future

§ Data security not included in the current definition, digital signatures

§ key component in .NET architecture

§ coming to J2EE 1.4 standard

§ also CA (Computer Associates) Unicenter for Web-services management system

§ platform-independent

§ XML could be slow when a real time service is needed between two systems

§ UDDI servers: MS, IBM, Ariba, NTT, Novell

Web Services example

Consumer

Directory UDDI

Send order

Confirm order; Delivery

Find service (WSDL)

Web site: shop

Electronic commerce

requesting checking

services

Service for checking address info

Service for checking credit info

XML in public administration

§ data transfer between systems:
 § police and magistrate
 § municipalities and central government
 § application forms
  • construction permits (one-stop-shopping)
  • unemployment benefits
  • agricultural subsidies and production reports
§ information dissemination services
 § weather conditions and forecasts
 § transport time tables
§ archives and storing of information
 § from microfiches to XML-data, the National Archive

RSS 2.0: feeds & podcasting

§ RSS 2.0 and Atom

§ information interchange on the Web

§ a way to syndicate blog posts and news sites

§ Amazon's OpenSearch technology uses RSS as a mechanism for providing search results and integrating search engines respectively

§ Podcasting to syndicate digital media content

§ Amazon's syndicated feeds

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<?xml-stylesheet title="XSL_formatting" type="text/xsl" href="/shared/bsp/xsl/rss/nolsol.xsl"?>

<rss version="2.0" xmlns:media="http://search.yahoo.com/mrss/">

<channel>
    <title>BBC News | News Front Page | World Edition</title>
    <link>http://news.bbc.co.uk/go/rss/-/2/hi/default.stm</link>
    <description>Visit BBC News for up-to-the-minute news, breaking news, video, audio and feature
    stories. </description>
    <language>en-gb</language>

    <lastBuildDate>Sun, 03 Feb 2008 09:57:26 GMT</lastBuildDate>
    <copyright>Copyright: (C) British Broadcasting Corporation</copyright>
    <docs>http://www.bbc.co.uk/syndication/</docs>
    <ttl>15</ttl>

<item>
    <title>Chad capital hit by new fighting</title>
    <description>Fresh fighting breaks out in  Chads capital NDjamena as rebels try for a second day
    to take control of the city.</description>
    <link>http://news.bbc.co.uk/go/rss/-/2/hi/africa/7224691.stm</link>
    <guid isPermaLink="false">http://news.bbc.co.uk/2/hi/africa/7224691.stm</guid>
    <pubDate>Sun, 03 Feb 2008 09:39:54 GMT</pubDate>
    <category>Africa</category>
    <media:thumbnail width="66" height="49"
    url="http://newsimg.bbc.co.uk/media/images/42523000/jpg/_42523051_rebels_index66_afp.jpg"/>
    </item>
    …
</channel>
</rss>
```

XML application areas:
XML is the basis for RDF and the Semantic Web

§ Resource Description Framework (RDF) is an XML text format that supports resource description and metadata applications

§ RDF integrates applications and agents into one Semantic Web

§ Formal descriptions of terms in a certain area (shopping or manufacturing, for example) are called ontologies

§ OWL-S version 1.0 (proposal) Ontology Web Language for services

§ CC/PP Composite Capability / Preference Profiles Structure and Vocabularies 1.0 specification for mobile devices

```xml
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-
    ns#" xmlns:s="http://example.edu/students/vocab#">
<rdf:Description rdf:about="http://example.edu/courses/6.001">
    <s:students>
    <rdf:Bag>
        <rdf:li rdf:resource="http://example.edu/students/Amy"/>
        <rdf:li rdf:resource="http://example.edu/students/Tim"/>
        <rdf:li rdf:resource="http://example.edu/students/John"/>
        <rdf:li rdf:resource="http://example.edu/students/Mary"/>
        <rdf:li rdf:resource="http://example.edu/students/Sue"/>
    </rdf:Bag>
  </s:students>
</rdf:Description>
</rdf:RDF>
```

Triples: subject-predicate-object expressions in RDF
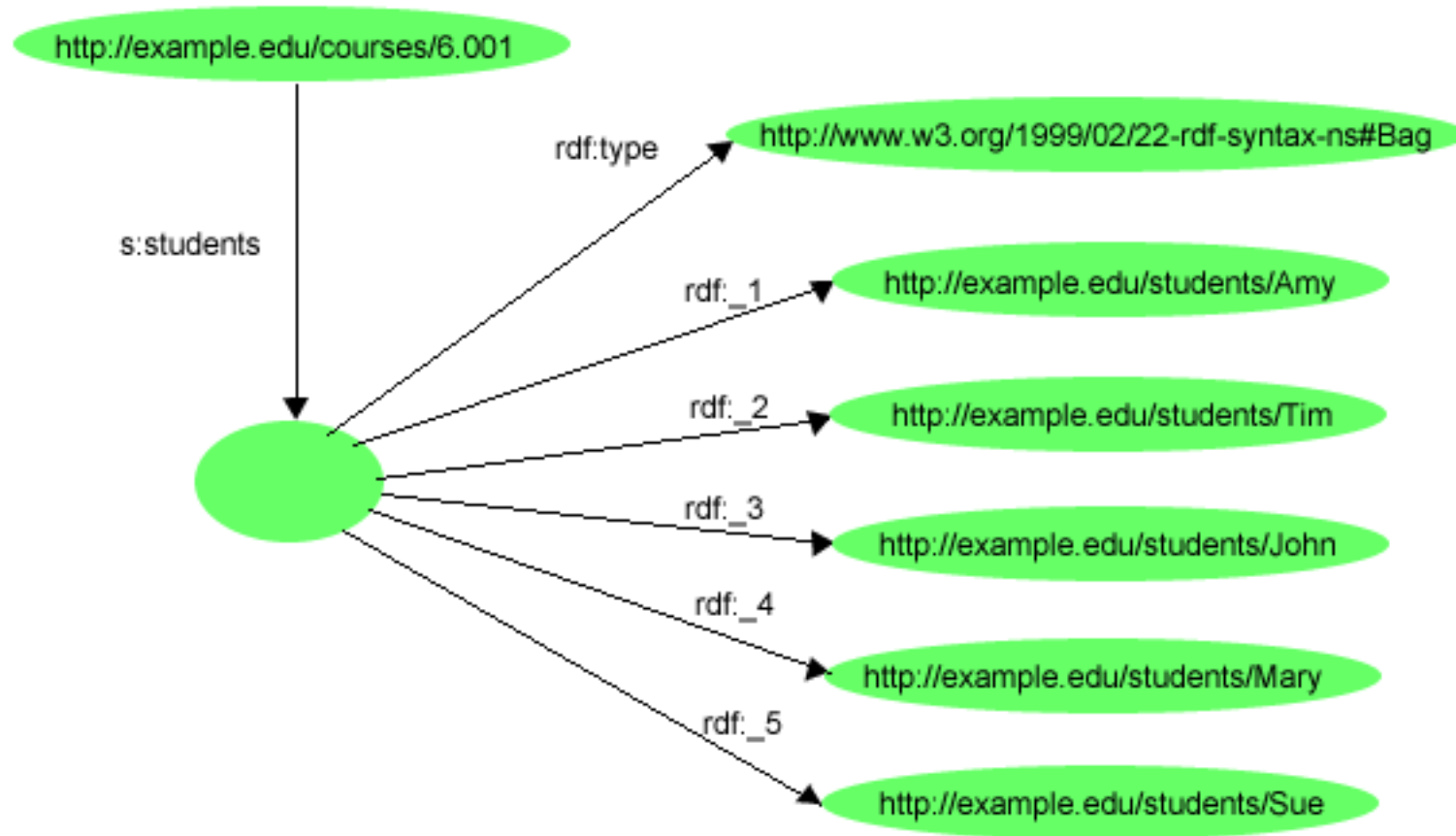
Recent metadata uses and standards

§ Open Linked Data: a recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF.

§ http://linkeddata.org/

§ http://data.gov.uk/

§ RDFa (or RDF in attributes) adds a set of attribute level extensions to XHTML for embedding rich metadata within Web documents.

§ Protocol for Web Description Resources (POWDER), a protocol for publishing descriptions of Web resources using RDF, OWL, and HTTP

§ Simple Knowledge Organization System (SKOS) for connection to structured vocabularies