



Konvertierung von Geschäftsdokumenten: Probleme und Lösungen mit XML/XSLT

Peter Buxmann, Erik Wüstner

Technische Universität Freiberg

Lehrstuhl für Wirtschaftsinformatik

<http://www.wiwi.tu-freiberg.de/wi>

<http://www.x-act.org>

Nürnberg, 10. September 2002

Agenda

1. Von EDI zu XML/EDI
2. Offene Standards und Open-Source für Austausch, Konvertierung und Integration von Geschäftsdokumenten
3. XML Konvertierungsprobleme / Vorstellung des Prototypen `<x:act>`
4. Ausblick

Nürnberg, 10. September 2002

Von EDI zu XML/EDI

- Der **Status Quo** von EDI: Killerapplikation und Erfolgsstory?
 - Vielzahl inkompatibler Standards
 - Häufig hohe Implementierungskosten
 - Zurzeit meist nur in großen Unternehmen
- Vision: Aufbau einer Lösung auf Basis offener Standards und Open-Source-Komponenten

Nürnberg, 10. September 2002

Offene Welt

- Offene Standards, z.B. XML, XSLT/XPath, XML Schema, http, SSL, Java, SOAP, WSDL
- Open-Source-Komponenten, z.B. Apache WebServer, Apache Xerces Parser, Apache Xalan XSLT-Processor, Apache Tomcat, Sun´s Java Web Service Developer Pack

Nürnberg, 10. September 2002

```
<?xml version="1.0" ?>
<?soxtype urn:x-commerceone:document:com:commerceone:XCBL30:XCBL30.sox$1.0?>
<Order>
  <OrderHeader>
    <OrderNumber>
      <BuyerOrderNumber>PO23540586</BuyerOrderNumber>
      <SellerOrderNumber>12949457S4</SellerOrderNumber>
      <ListOfMessageID>
        <MessageID>
          <IDNumber>Signed Purchase Order 12335</IDNumber>
          <IDAssignedBy>
            <IDAssignedByCoded>Other</IDAssignedByCoded>
            <IDAssignedByCodedOther>LaMu2386</IDAssignedByCodedOther>
          </IDAssignedBy>
          <IDAssignedDate>20010211T15:25:00</IDAssignedDate>
        </MessageID>
      </ListOfMessageID>
    </OrderNumber>
    <OrderIssueDate>20010225T09:30:00</OrderIssueDate>
    <OrderReferences>
      <AccountCode>
        <Reference>
          <RefNum>58395776904</RefNum>
          <RefDate>20010211T15:26:00</RefDate>
        </Reference>
      </AccountCode>
    </OrderReferences>
  </OrderHeader>
</Order>
```

XML-Standardisierungsinitiativen:

- xCBL
- CEN/ISSS XML-EDI
- DIN 16557-4
- XML/EDI Group
- cXML
- OAGIS
- ...



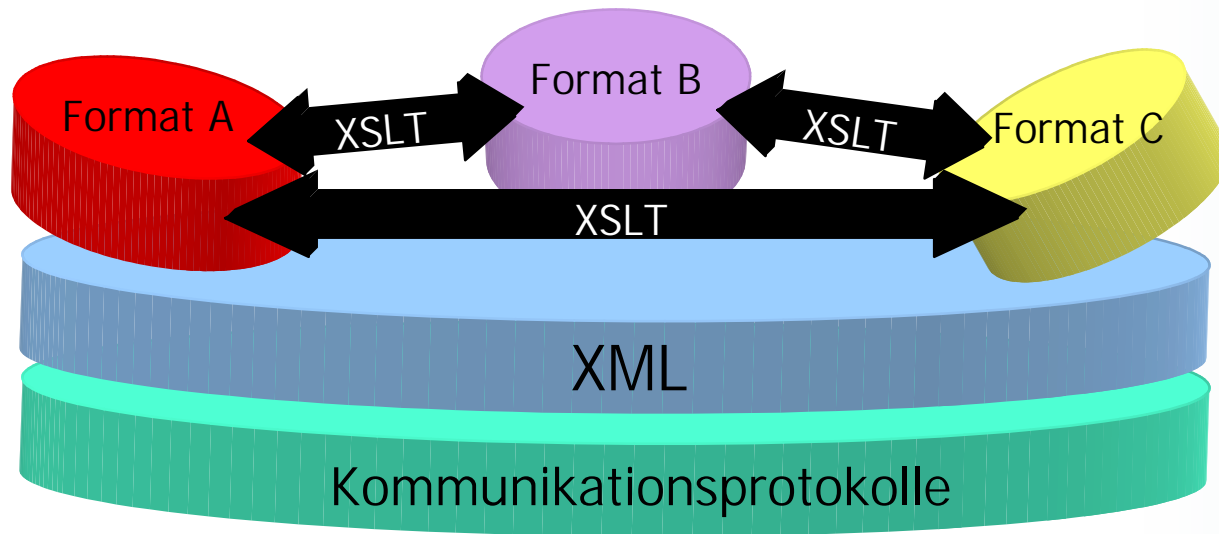
basieren auf EDIFACT-
bzw. X12-Strukturen

XML-Frameworks:

- BizTalk
- ebXML
- RosettaNet
- ...

Unser Konvertierungsansatz

Schichtenmodell: Die Beteiligten einigen sich auf Basisstandards in der unteren Ebene, können aber auf höherer Ebene ein Format frei wählen.



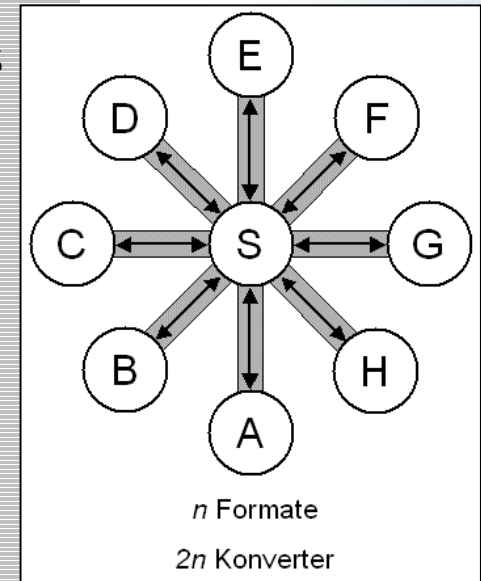
Nürnberg, 10. September 2002

Rahmenbedingungen

- Konvertierung durch Nutzung eines mittelnden Formats („Superstandard“)
- Dokumente sind in XML beschrieben und es liegt ein entsprechendes XML-Schema vor
- Vorerst kein Rückgriff auf Ontologien bzw. RDF
- Detaillierte Kenntnis der beteiligten Formate

Zielsetzung

- Any-to-Any-Konvertierung
- Offene und wieder verwendbare Konverter (XSLT)
- Umfassende graphische Unterstützung bei der Stylesheet-Erstellung
- Verringerung des Mapping-Aufwandes, da automatisch beide Stylesheets erstellt werden.



1. Das Markup ist verschieden, der semantische Gesamtgehalt ist gleich.

a)

```
<Person>
  <FirstName>John</FirstName>
  <LastName>Doe</LastName>
</Person>
```

b)

```
<PersonalData>
  <Firstname>John</Firstname>
  <FamilyName>Doe</FamilyName>
</PersonalData>
```

```
<xsl:element name="PersonalData">
  <xsl:element name="Firstname"><xsl:value-of
select="/Person/FirstName"/></xsl:element>
  <xsl:element name="FamilyName"><xsl:value-of
select="/Person/LastName"/></xsl:element>
</xsl:element>
```

2. Unterschiedliches Core-Data-Format, bei gleicher Core-Data.

a)

```
<Date>11/15/2001</Date>
```

b)

```
<Date>15.11.2001</Date>
```

```
<xsl:variable name="Date" select="./Date"/>
<xsl:element name="Date">
  <xsl:value-of select=
    "concat(substring($Date,3,2),'.',substring($Date,1,2),'.',
    substring($Date,7,4))"/>
</xsl:element>
```

3. Unterschiedliche Strukturtiefe, wobei die Struktur selbst Meta-Informationen enthält.

a)

```
<Person>  
  <FirstName>John</FirstName>  
  <LastName>Doe</LastName>  
</Person>
```

b)

```
<PersonalData>  
  <Name>John Doe</Name>  
</PersonalData>
```

```
<xsl:variable name="FirstName" select="/Person/FirstName"/>  
<xsl:variable name="LastName" select="/Person/LastName"/>  
<xsl:element name="Name">  
<xsl:value-of select="concat($FirstName, ' ', $LastName)"/>  
</xsl:element>
```

4. Gleiche Core-Data, unterschiedlicher semantischer Gesamtgehalt.

<p>a)</p> <pre><PhoneContact> <PhoneNr>1-800-37342374</PhoneNr> <PhoneNr>917-816-9494</PhoneNr> </PhoneContact></pre>	<p>b)</p> <pre><PhoneContact> <FixedNetwork> <PhoneNr>1-800-37342374</PhoneNr> </FixedNetwork> <CellularNetwork> <PhoneNr>917-816-9494</PhoneNr> </CellularNetwork> </PhoneContact></pre>
---	---

5. Redundante Daten.

<p>a)</p> <pre><InvoiceDocument> <ItemList> <Item> <ItemNr>001</ItemNr> <Name>Product 1</Name> <NetPrice>100.00</NetPrice> <VATRate>10.00</VATRate> </Item> - <Item> <ItemNr>002</ItemNr> <Name>Product 2</Name> <NetPrice>200.00</NetPrice> <VATRate>10.00</VATRate> </Item> </ItemList> </InvoiceDocument></pre>	<p>b)</p> <pre><Invoice> - <Items> - <Item> <ItemCode>001</ItemCode> <ItemName>Product 1</ItemName> <ItemNetPrice>100.00</ItemNetPrice> <VAT>10.00</VAT> </Item> - <Item> <ItemCode>002</ItemCode> <ItemName>Product 2</ItemName> <ItemNetPrice>200.00</ItemNetPrice> <VAT>10.00</VAT> </Item> </Items> <NetTotal>300</NetTotal> <VATTotal>30</VATTotal> <GrossTotal>330</GrossTotal> </Invoice></pre>
---	---

```
<xsl:element name="NetTotal">
  <xsl:value-of select="sum(//NetPrice)"/>
</xsl:element>
```

berg, 10. September 2002

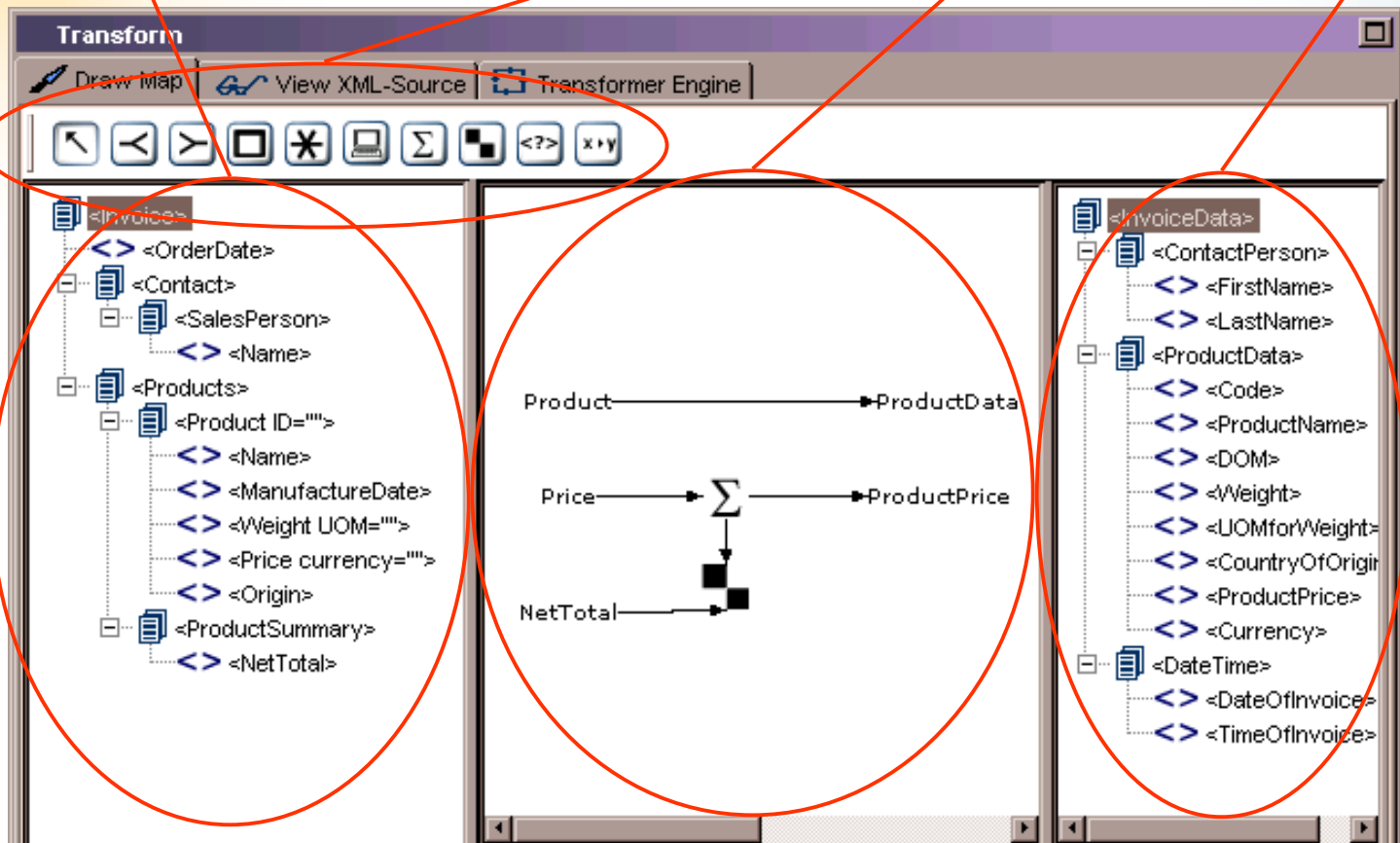
Lösung der Probleme in <x:act>

Schema des Quelldokuments

Tools

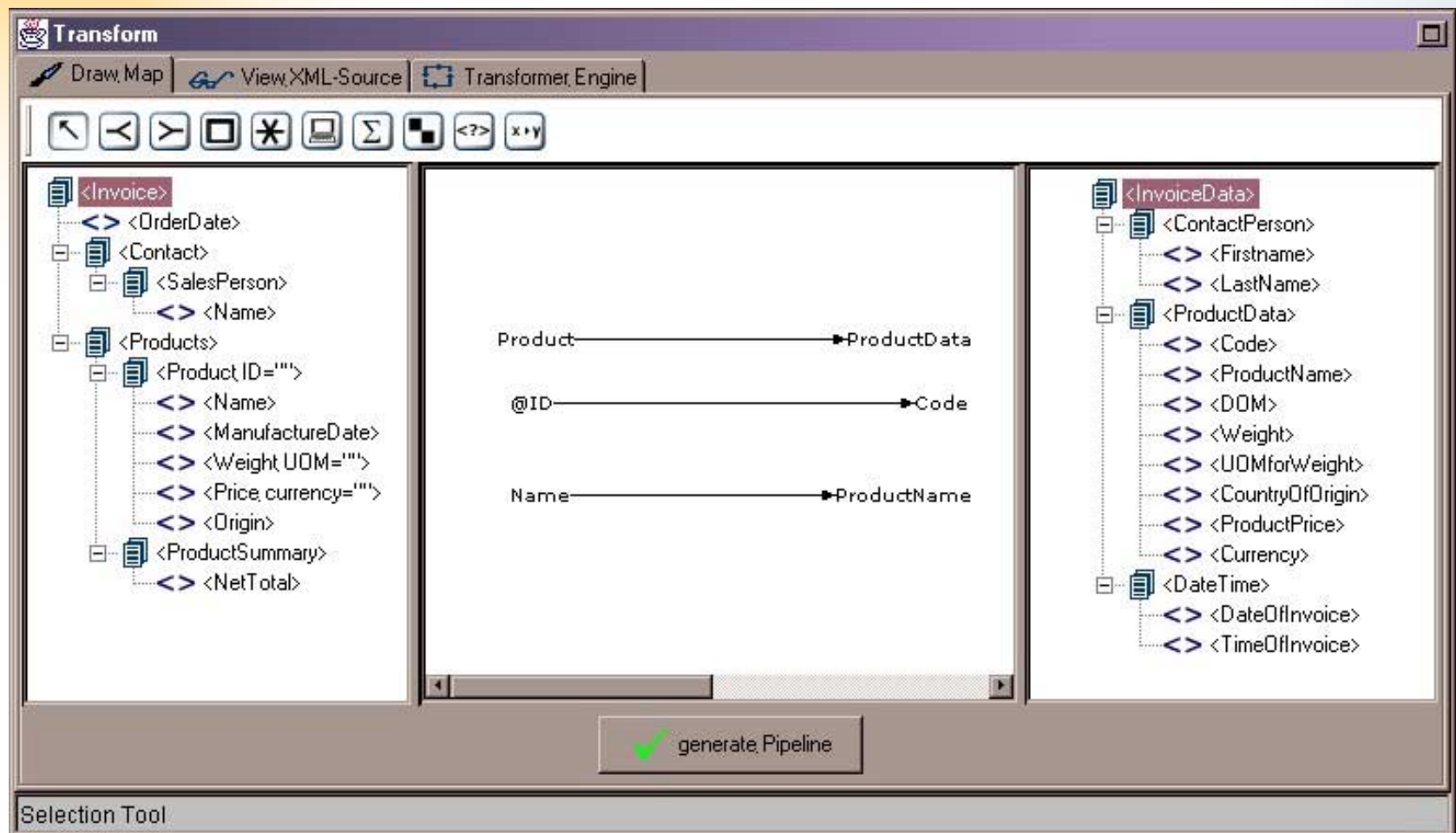
Mapping-Bereich

Super-Standard



Nürnberg, 10. September 2002

1. Das Markup ist verschieden, der semantische Gesamtgehalt ist gleich



The screenshot shows the Transform software interface with three main panes:

- Left Pane (Source XML):** A tree view of an `<Invoice>` document. It contains:
 - `<OrderDate>`
 - `<Contact>`
 - `<SalesPerson>`
 - `<Name>`
 - `<Products>`
 - `<Product ID='''>`
 - `<Name>`
 - `<ManufactureDate>`
 - `<Weight, UOM='''>`
 - `<Price, currency='''>`
 - `<Origin>`
 - `<ProductSummary>`
 - `<NetTotal>`

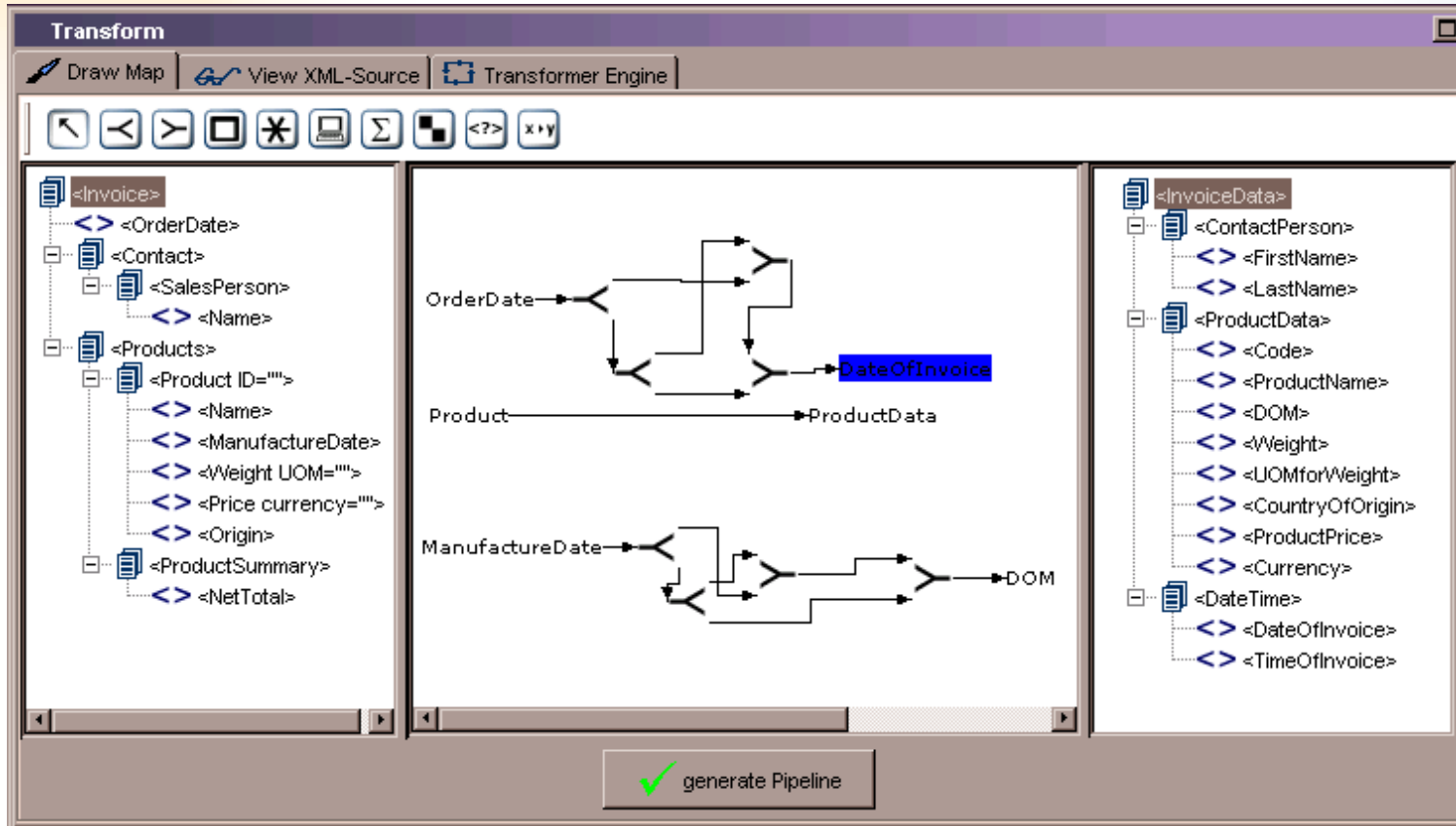
- Middle Pane (Transformation Rules):** A diagram showing mappings:
- `Product` → `ProductData`
- `@ID` → `Code`
- `Name` → `ProductName`
- Right Pane (Target XML):** A tree view of an `<InvoiceData>` document. It contains:
- `<ContactPerson>`
 - `<Firstname>`
 - `<Lastname>`
- `<ProductData>`
 - `<Code>`
 - `<ProductName>`
 - `<DOM>`
 - `<Weight>`
 - `<UOMforWeight>`
 - `<CountryOfOrigin>`
 - `<ProductPrice>`
 - `<Currency>`
- `<DateTime>`
 - `<DateOfInvoice>`
 - `<TimeOfInvoice>`

At the bottom of the interface is a button labeled "generate Pipeline" with a green checkmark icon.

Nürnberg, 10. September 2002

2. Unterschiedliches Core-Data-Format, bei gleicher Core-Data

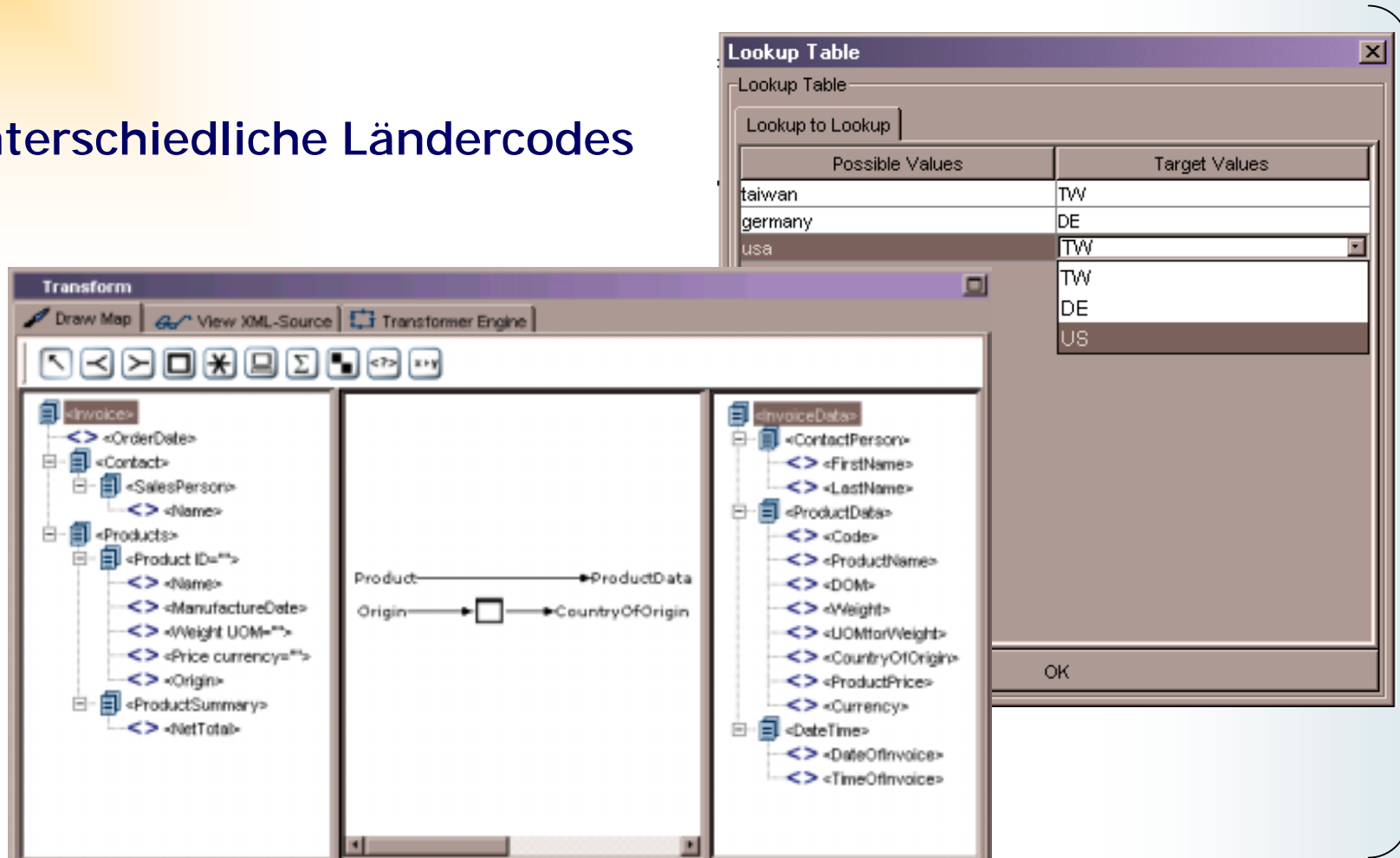
Unterschiedliche Datumsformate



Nürnberg, 10. September 2002

2. Unterschiedliches Core-Data-Format, bei gleicher Core-Data

Unterschiedliche Ländercodes



The screenshot shows the 'Transform' tool interface with a mapping diagram and a 'Lookup Table' dialog. The mapping diagram shows the following structure:

```

    graph LR
      Product --> ProductData
      Origin --> CountryOfOrigin
  
```

The 'Lookup Table' dialog is titled 'Lookup Table' and contains a table with the following data:

Possible Values	Target Values
taiwan	TW
germany	DE
usa	TW
	TW
	DE
	US

The 'Transform' tool shows the following XML structures:

Source XML:

```

    <?xml version="1.0" encoding="UTF-8" ?>
    <invoice>
      <OrderDate>
      <Contact>
        <SalesPerson>
          <Name>
        <Products>
          <Product ID="">
            <Name>
            <ManufactureDate>
            <Weight UOM="">
            <Price currency="">
            <Origin>
          <ProductSummary>
            <NetTotal>
    
```

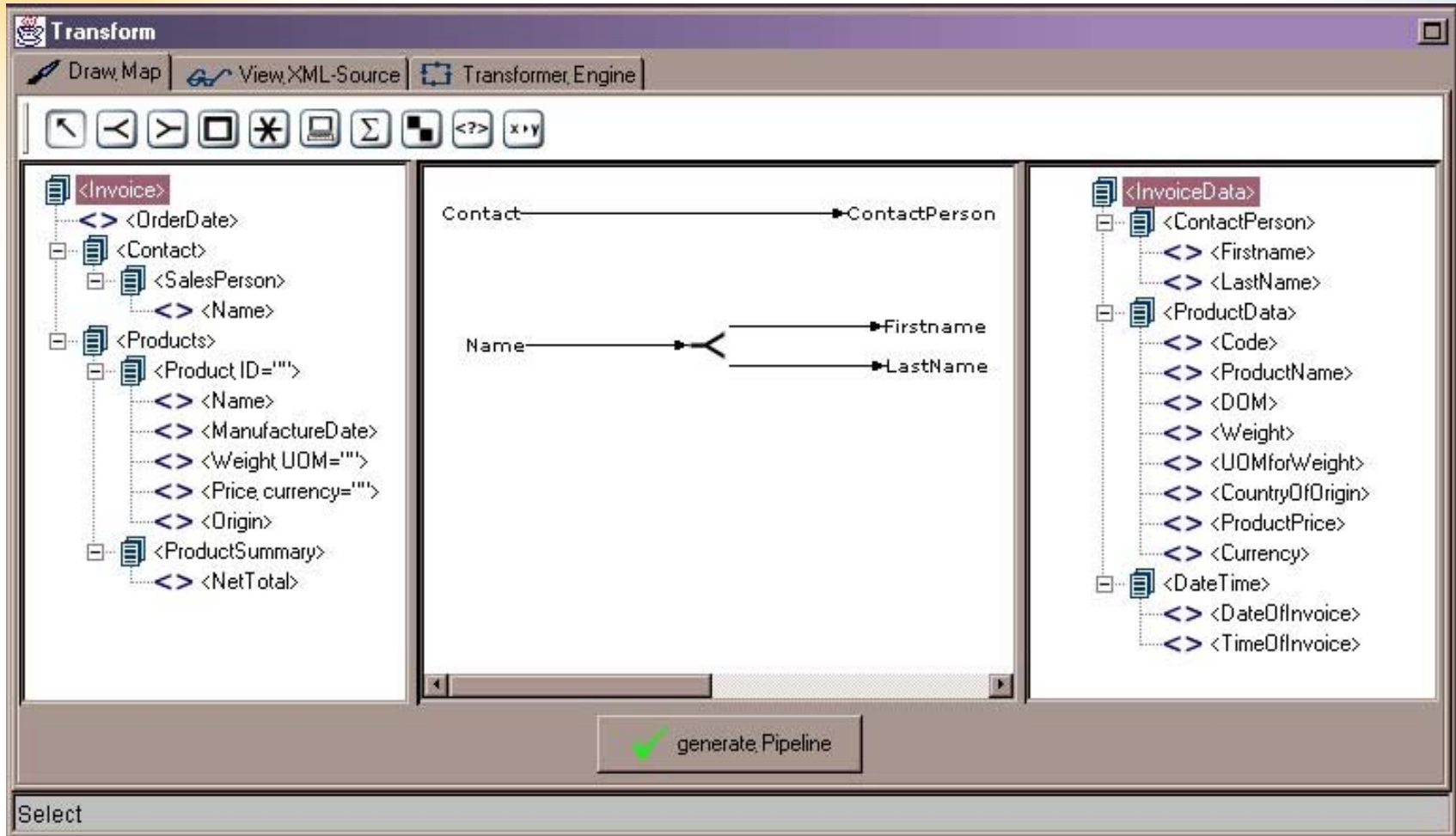
Target XML:

```

    <?xml version="1.0" encoding="UTF-8" ?>
    <invoiceData>
      <ContactPerson>
        <FirstName>
        <LastName>
      <ProductData>
        <Code>
        <ProductName>
        <DOM>
        <Weight>
        <UOMforWeight>
        <CountryOfOrigin>
        <ProductPrice>
        <Currency>
      <DateTime>
        <DateOfInvoice>
        <TimeOfInvoice>
    
```



Nürnberg, 10. September 2002

3. Unterschiedliche Strukturtiefe, wobei die Struktur selbst Meta-Informationen enthält

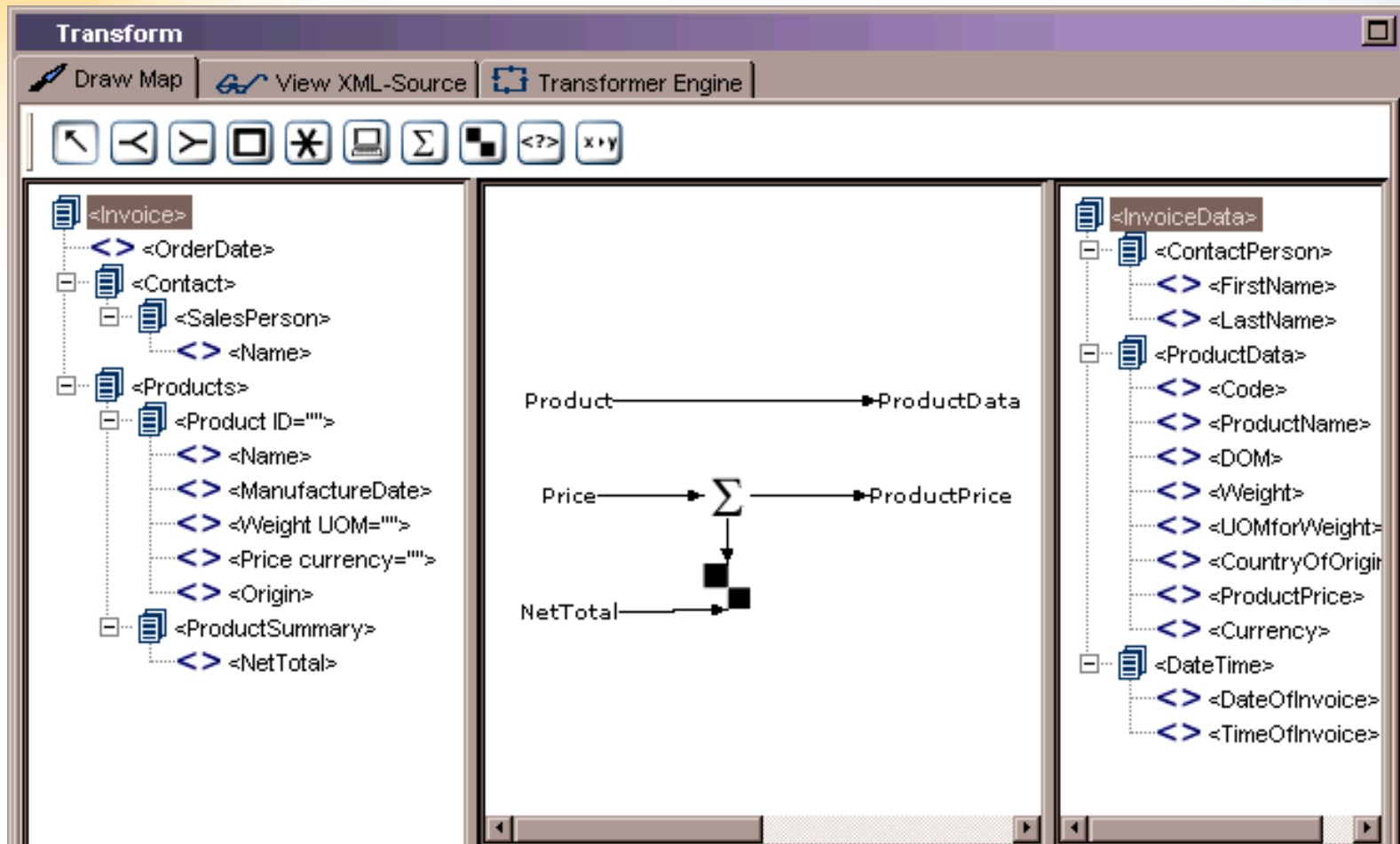


Nürnberg, 10. September 2002

4. Gleiche Core-Data, unterschiedlicher semantischer Gesamtgehalt

- Genaue Konvertierung nicht möglich
- In der Praxis Konvertierung möglich, aber unzuverlässig
- Einsatz der Tools:
 - Multiplier A square icon with a black border containing a white asterisk symbol.
 - Extractor A square icon with a black border containing a white XML-like tag symbol '<?>'.

5. Redundante Daten



Nürnberg, 10. September 2002



- Aufbau eines Web Service zur Konvertierung von XML-Dokumenten, erster Prototyp und weitere Informationen unter www.x-act.org verfügbar
- Fragen, Hinweise, Anregungen?
 - buxmann@bwl.tu-freiberg.de
 - erik.wuestner@bwl.tu-freiberg.de

Nürnberg, 10. September 2002