

XHTML

www.skillspartner.com

Skillspartner

We turn information into skills

Mål

Deltagarna ska förstå:

- Vad XML är
- Hur XML fungerar
- Vilka problem XML löser
- Hur XML används
- Känna till begrepp: DTD, DOM, XSL etc.
- Känna till verktyg

Innehåll

- Vad är XML
- XML i relation till HTML
- XML – Terminologi
- Verktyg för t.ex. editering av XML

Lärare

Mats Strandberg

Cognitudo AB

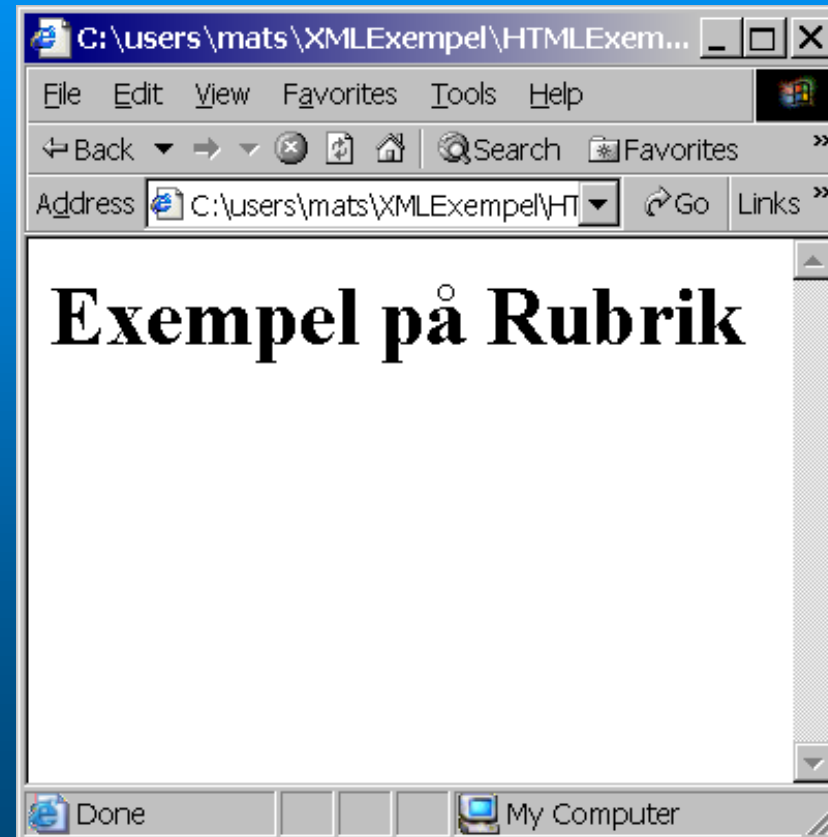
mats@cognitudo.com

Vad är XML? ⁽¹⁾

- XML är ett märkordsspråk
- HTML är ett annat märkordsspråk

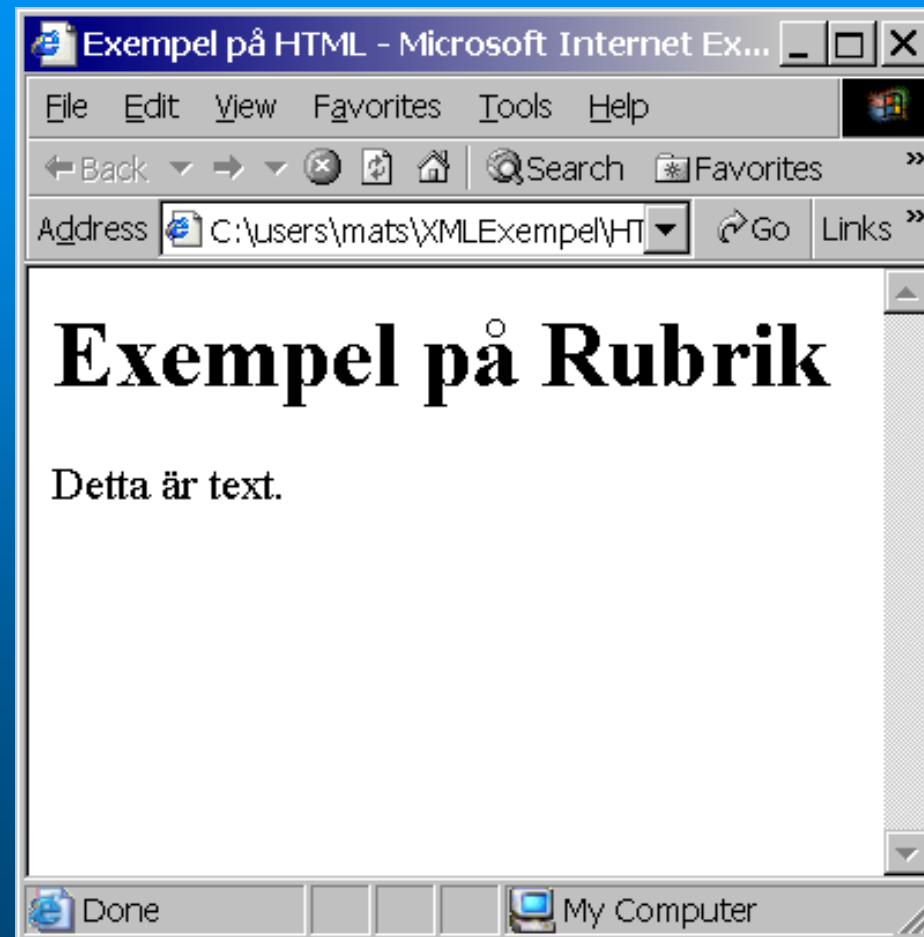
HTML Exempel 1

```
<html>
  <body>
    <h1>
      Exempel på Rubrik
    </h1>
  </body>
</html>
```



HTML Exempel 2

```
<html>
<head>
  <title>
    Exempel på HTML
  </title>
</head>
<body>
  <h1>
    Exempel på Rubrik
  </h1>
  <p>
    Detta är text.
  </p>
</body>
</html>
```



HTML Telefonlista

```
<html>
  <head>
    <title>Telefonlista</title>
  </head>
  <body>
    <h1>Telefonlista</h1>
    <b>Gustavsson</b>
    <em>Agda</em>
    <code>01234567</code>
    <br>
    <b>Larsson</b>
    <em>Malte</em>
    <code>01234321</code>
  </body>
</html>
```



HTML Telefonlista v2

```
<html>
  <head>
    <title>Telefonlista v2</title>
  </head>
  <body>
    <h1>Telefonlista v2</h1>
    Agda Gustavsson
    <b>0706429326</b><br>
    Malte Larsson
    <b>01234321</b>
  </body>
</html>
```



HTML för text och data

- På webben finns textdokument i HTML och data i HTML
- HTML är bra för att *presentera* text, data, bilder och hyperlänkar etc.
- HTML är *inte* bra för att *överföra* data
- Screen scraping är inte en stabil lösning!

Representation av info

```
<Telelista>
  <Person>
    <Förnamn>Agda</Förnamn>
    <Efternamn>Gustavsson</Efternamn>
    <Telnr>01234567</Telnr>
  </Person>
  <Person>
    <Förnamn>Malte</Förnamn>
    <Efternamn>Larsson</Efternamn>
    <Telnr>01234321</Telnr>
  </Person>
</Telelista>
```



Vad är XML? ⁽³⁾

- XML – eXtensible Markup Language
- XML används för att skapa egna märkordsspråk!

Definiera märkordsspråk

- Ett nytt märkordsspråk definieras i en DTD
- DTD – Document Type Definition
- DTD Beskriver *vilka* taggar som får förekomma
- DTD Beskriver *hur* taggar får kombineras

DTD för telefonbok

```
<?xml version="1.0" encoding="ISO-8859-1"?>  
<!ELEMENT Telelista (Person*)>  
<!ELEMENT Person (Förnamn, Efternamn, Telnr)>  
<!ELEMENT Förnamn (#PCDATA)>  
<!ELEMENT Efternamn (#PCDATA)>  
<!ELEMENT Telnr (#PCDATA)>
```

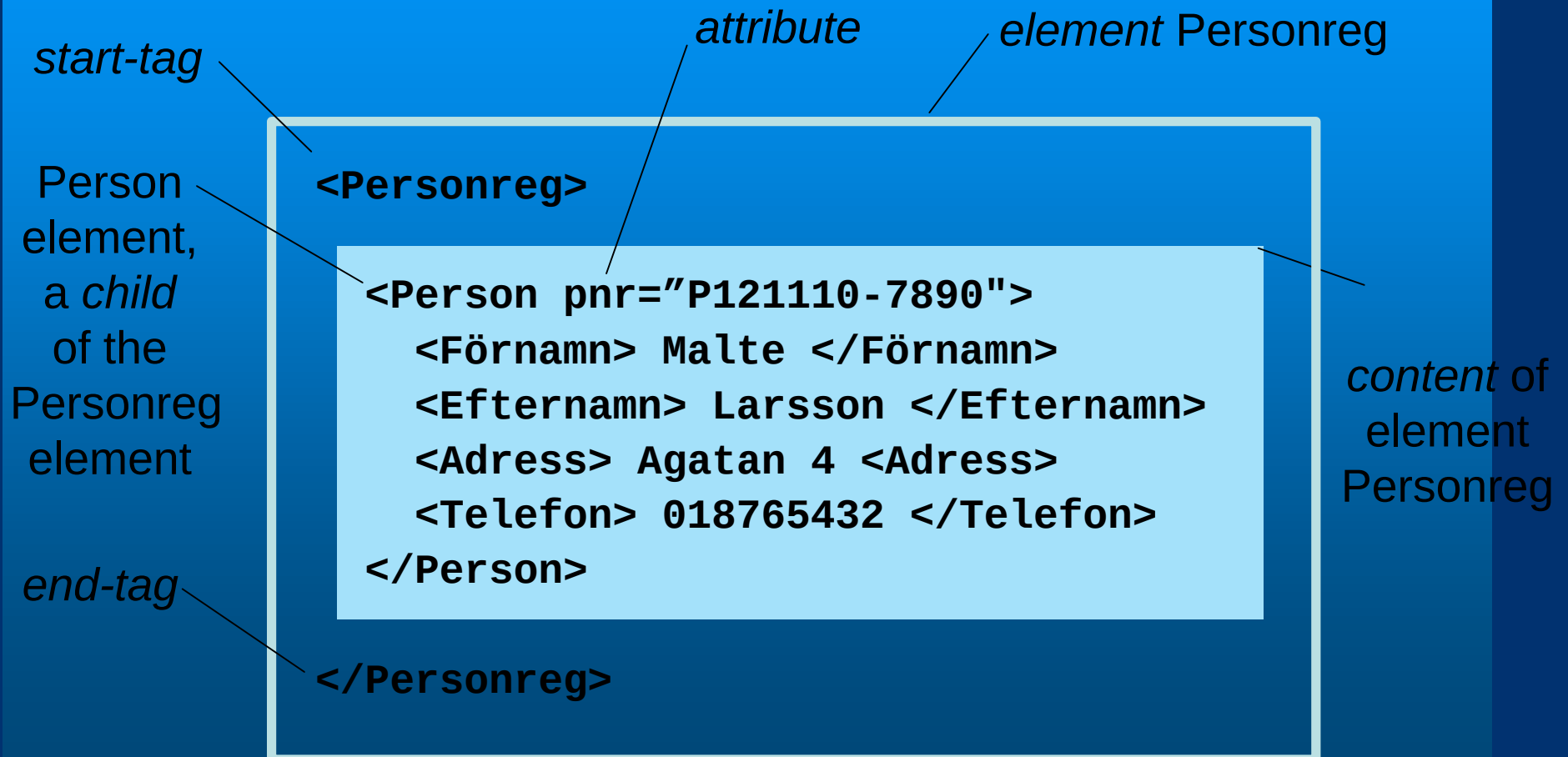
Komplett XML-fil för telbok

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE Telelista SYSTEM "telelista.dtd">
<Telelista>
  <Person>
    <Förnamn>Agda</Förnamn>
    <Efternamn>Gustavsson</Efternamn>
    <Telnr>01234567</Telnr>
  </Person>
  <Person>
    <Förnamn>Malte</Förnamn>
    <Efternamn>Larsson</Efternamn>
    <Telnr>01234321</Telnr>
  </Person>
</Telelista>
```

Vad är XML? ⁽³⁾

- XML – eXtensible Markup Language
- XML används för att definiera märkordsspråk!
- XML är en standard som förvaltas av W3C (World Wide Web Consortium)

XML – Termer



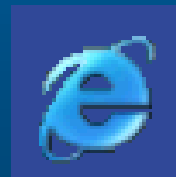
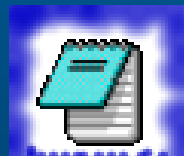
XML och DTD

- Regler för märkordsspråk beskrivs av XML 1.0
 - En <Tag> måste ha en motsvarande </Tag> etc.
- DTD beskriver strukturella regler
 - T.ex. En <Person> kan innehålla ett <Förnamn> men inte ytterligare en <Person>
- Validering av XML-dokument
 - Välformade (Eng: well-formed) XML-dokument – uppfyller XML 1.0
 - Giltiga (Eng: Valid) XML-dokument – följer DTD

Verktyg för XML

Exempel på verktyg för editering av XML:

- Notepad + Internet Explorer
- XMLSpy
- Etc.



XML Bakgrund

- Idéer från SGML (ISO 8879)
- HTML (Tim Berners-Lee. Idéer från SGML)
- Version av XML:
 - XML 1.0 (2ed) 2000-10
 - XML 1.0 (1ed) 1998-02

Vad är XML?

- XML – eXtensible Markup Language
- Ett sätt att lagra data i textformat
- W3C Standard för att ”märka” data
- Ett stöd för att beskriva format för data som skall utväxlas mellan olika datorsystem

Vad löses med XML

- Format som är läsbart för människor
- Lagring av dokument på ett "uppgraderingssäkert" format.
- Utbyte av information mellan system
- Skiljer innehåll från presentation
- En källa, många vyer

XML läsbart för människor

Detta innebär:

- XML-dokument innehåller *enbart* text i en namngiven teckenuppsättning, ASCII, ISO8859-1 etc.
- Jämför med binärformat (t.ex. databas eller word-dokument)
- Kan editeras i "alla" editorer
- Portabelt mellan OS, Processorer etc.

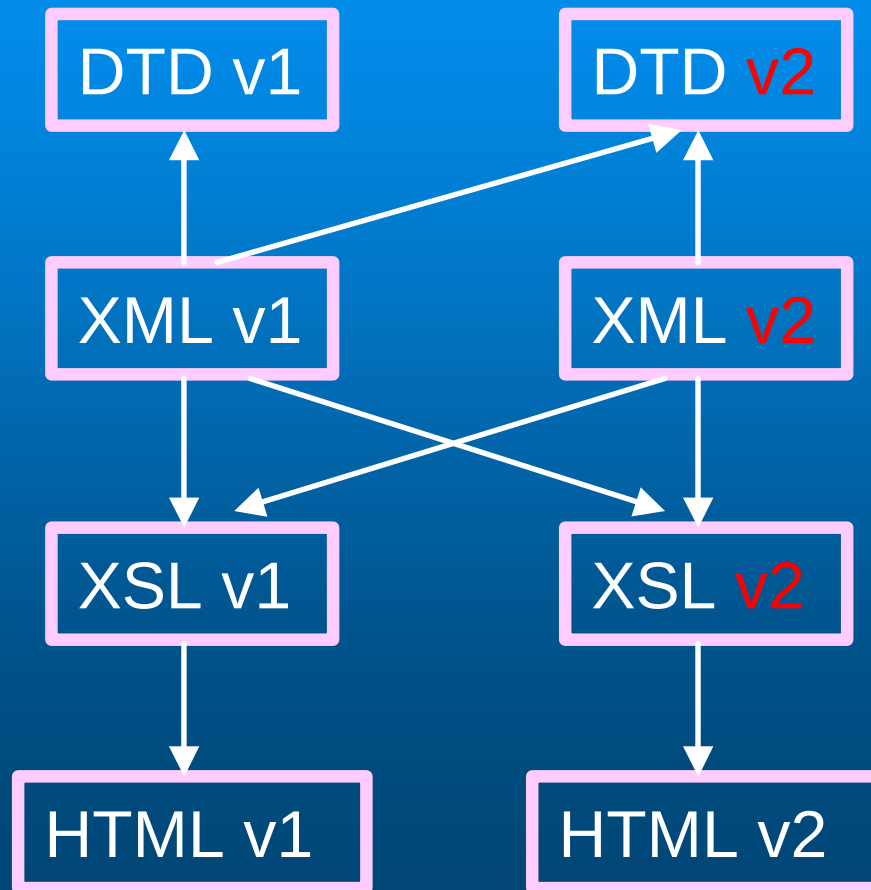
”Uppgraderingssäkert” format

- Taggar används istället för offset etc.
- Innehåll i nya taggar kan ignoreras av gamla tillämpningar.
- Taggar kan göras optionella (svårare med binärformat)
- Format kan versionshanteras

Utbyte av data mellan system

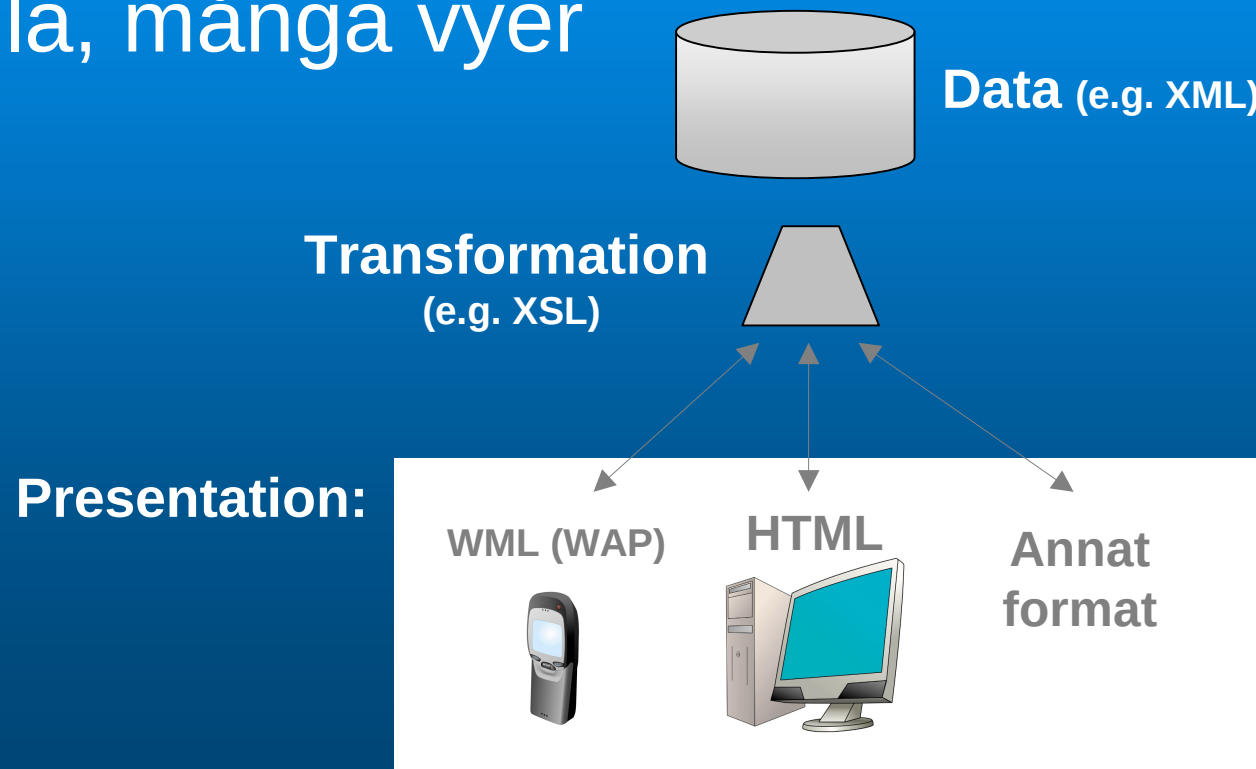
- XML är text - plattformsoberoende
- Ramverk för att definiera syntax (men ej semantik). Vad betyder t.ex. `<a><b/><a>`?
- Bakåt kompatibla förändringar i formatet kan införas "osynkroniserat"

Versionsberoende



Innehåll skilt från presentation

- Källinformation åtskilt från presentation
- En källa, många vyer



Tillämpningar av XML

- XHTML eXtensible Hypertext Markup Language
- WML (WAP) Wireless Markup Language
- VML Vector Markup Language
- VoiceXML
- XUL Xml-based User interface Language
- XBRL eXtensible Business Reporting Language
- RDF Resource Description Framework
- Etc.

Användningsområden

- Beskrivningsfiler (VML, RDF)
- Presentationsformat (XHTML, WAP)
- Deployment descriptors
- Konfigurationsfiler
- Logfiler
- Make-filer (ANT)
- Programmeringsspråk (VoiceXML)

DOM

- DOM – Document Object Model
- Trädstruktur för hantering av XML-data från programkod (Java, C++ etc.)
- Flexibel, men komplex, struktur
 - Alternativ JAXB
- Hela trädet i primärminnet
 - Alternativ SAX

XMLSchema

- Alternativ till DTD
- Hårdare specifikation av datatyper än DTD
- T.ex. Datum-typ, URL-typ
- XML Schema 1.0 – 2001-05

XML eller Databas?

XML sekventiell data. Databas via nyckel etc. DB har transaktion.

	DB	XML
ETT dokument uppdateras av EN användare		X
Små till medelstora datamängder utan uppdatering		X
Små datamängder med uppdatering	X	X
Flera simultiga användare	X	(X)
Stora datamängder	X	
Snabb åtkomst	X	

XPath

- Icke XML
- Refererar till delar av XML dokument
- Används av XSLT
- Exempel på Xpath i XSLT:

```
<xsl:for-each select="Telelista/Person">  
<xsl:value-of select="Telelista/@version"/>
```

XPointer

- Icke XML, Kan bara peka på XML-dok
- Utökning av Xpath (med "point" och "ranges")
- Exempel:

```
xpointer(/Personreg/Person[@id<2000000000000])
```

- Används som "fragment i URL":

```
http://cognitude.com#xpointer(/Personreg/Person[@id<2000000000000])
```

XLink

- XML Syntax för länkar
- Enkelriktade, dubbelriktade, en-till-många
- Kan användas i webbrowser eller av XSLT
- Syntax men ej semantik

```
<Person xlink:type="simple"
        xlink:title="CV för Anders Andersson"
        xlink:href="www.anders.se/cv"
        xlink:actuate="onLoad" xlink:show="embed">
```

Anders Andersson

```
</Person>
```

XML URLar

- w3c.org
- microsoft.com/xml (sök-url)
- sun.com/xml (sök-url)
- xml.apache.org
- www.cognititude.com/xml

Sammanfattning

- Textformat (med definierad teckenuppsättn.)
- XML skiljer data från presentation
- Lätt att skapa väldefinierat format (DTD, XML Schema)
- Utökbart = mindre behov av versionsynk.
- Etablerad W3C standard =
verktygsstöd, enkel kommunikation etc.