



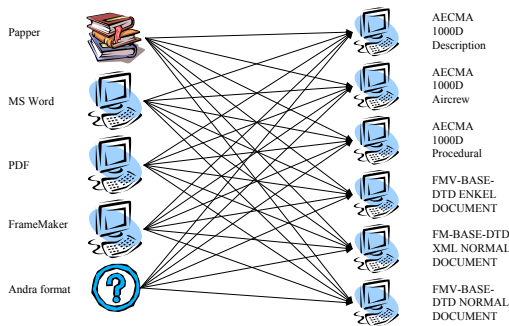
## XML-Schema för validering av skrivregler vid modernisering av informationsarvet

2004-05-18  
Joakim Ström, Projektledare

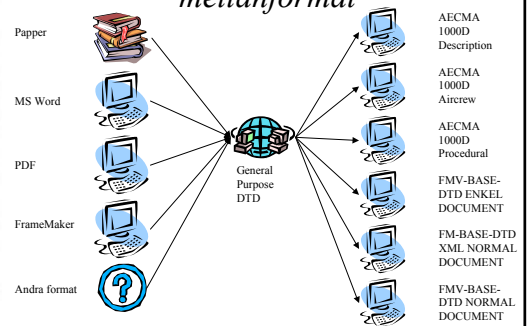
## Uppdraget att modernisera FMVs informationsarv

- Behov av elektronisk dokumentation
- Många källformat
- Ostrukturerad information
- Många olika DTDer
- Upphandling av konverteringstjänst

## Konverteringsbehov



## Förenklad konvertering med mellanformat



## Produktionsformat

Ett enhetligt produktionsformat. Krav:

- ✓ Funktionell produktionsmiljö
- ✓ Lätt att förstå och använda
- ✓ Enkelt underhåll
- ✓ Informationsanalys och validering
- ✓ Möjlig att publicera och konvertera
- ✓ Skall kunna lagra alla befintliga dokument

## Leveransformat

Behovsstyrd leverans till många heterogena leveransformat. Exempel:

- Specifika XML/SGML-DTDer
- HTML
- PDF
- Papper
- MS Word

## Leverans

- Specifik valideringsrutin för varje leveransformat (XPath-baserad)
- Olika utskriftsformat via XSL-FO
- XSLT-publicering till HTML
- XSLT-konvertering till andra DTDer för export till tekniska system
- Utskrifts- och publiceringsrutiner från målformaten kan elimineras

## Process för konvertering



## Validering i två steg

- Minimal DTD bestämmer formatets grammatik. Automatisk validering vid redigering.
- Fallspecifika affärsregler bestämmer formatets semantik. Validering vid behov.

## General Purpose DTD

- Mycket tillåtande regler
- Samma byggstenar används överallt
- Litet antal generiska element definierar strukturregler (jfr ordklasser)
- Attribut definierar elementets betydelse (jfr ord)

## Exempel

	Element (Ordklass)	Attribut (Ord)
<pre>&lt;wrapper class="note"&gt;   &lt;paragraph&gt;Leveranstid tio   arbetsdagar&lt;/paragraph&gt; &lt;/wrapper&gt;</pre>	wrapper	note
<pre>&lt;meta-data&gt;   &lt;meta-block class="date"&gt;     &lt;meta-text     class="year"&gt;2003&lt;/meta-text&gt;     &lt;meta-text     class="month"&gt;06&lt;/meta-text&gt;     &lt;meta-text     class="day"&gt;06&lt;/meta-text&gt;   &lt;/meta-block&gt; &lt;/meta-data&gt;</pre>	meta-block	date
	meta-text	year
	meta-text	month
	meta-text	day

## Varför inte General Purpose Schema?





## *Skrivregler*

- Definierar en specifik vokabulär
- Olika skrivregler kan användas i olika sammanhang
- Validering med XML Schema
- Kompletteras med förklarande anvisningar



## *Utgångspunkter*

- Handbok FMV Materielpublikationer
- Befintliga publikationer
- Mål-DTDer

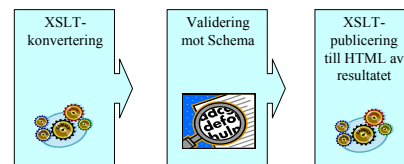


## *XML Schema användning*

- Definierar en delmängd av DTDn
- Vokabulären kan begränsa strukturen
- XSLT-försteg till valideringen gör unika element av alla ”ord”



## *Valideringsprocess*



## *Designprinciper*

- DTD-liknande struktur
- Lokala elementdeklarationer
- Element- och attributgrupper
- Typdefinitioner
- ”Mixed content”
- Inget namespace



## *Vidare perspektiv*

- General Purpose med ett specifikt syfte?
- Nyproduktion av publikationer