

XML Schema

David Massey


MBIB4140

29-8-2017

Structured information toolkit

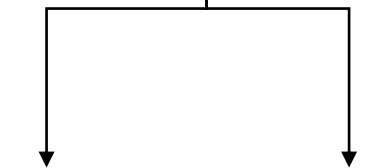
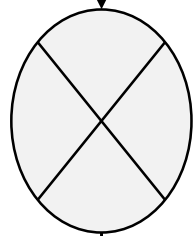
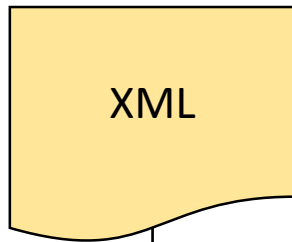
"XML's new playmates include stylesheets for display and transformation, strong methods for linking resources, tools for data manipulation and querying, **error checking and structure enforcement** tools, and a plethora of development environments. As a result of these applications, XML is assured a long and fruitful career as the structured information toolkit of choice."

Ray, Learning XML, 2003, s. xi



Uke	Forleser	Dato	Forelesning tittel	Pensum
1	David	2017-08-22	Introduksjon. XML og databaser, en oppfriskning	
2	David	2017-08-29	XML Schema	W3C - XSD
3	Thomas	2017-09-05	XPath og XQuery	W3C - XQuery, XPath
4	David	2017-09-12	XSLT 1	Tidwell, W3C - XSLT
5	David	2017-09-21	XSLT 2	Tidwell, W3C - XSLT
6	David	2017-09-26	XML-stacken workshop	

Velformet?

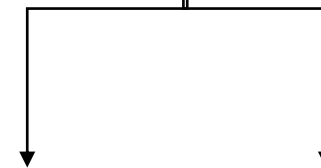
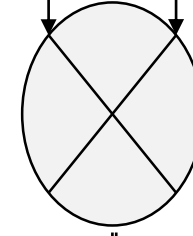
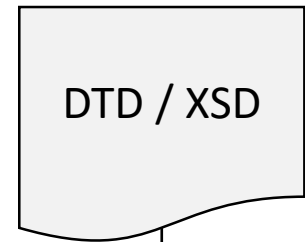
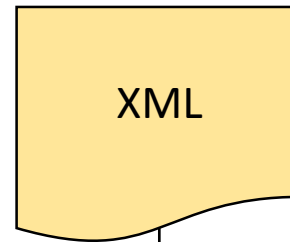


Ja

Nei

Program som kan
sjekke om et XML
dokument er
velformet

Gyldig?



Ja

Nei

Program som kan
sjekke om et XML
dokument er gyldig

DTD eller XSD

- DTD = Document Type Definition
 - Henger igjen fra SGML
 - Spesielt for enkelt innhold
 - Fortsatt i bruk pga eldre løsninger i systemer
- XSD = XML Schema Definition
 - Basert på XML
 - Kraftigere
 - Støtter navnerom og datatyper

Gyldig (valid)

XML Schemas express shared vocabularies and allow machines to carry out rules made by people. They provide a means for defining the structure, content and semantics of XML documents.

<https://www.w3.org/XML/Schema>

XML skjema definerer

- Elementer og attributter i et XML dokument
- Rekkefølgen av elementer
- Hvilke elementer er underordnede elementer
- Antall underordnede elementer
- Om elementer kan være tom, eller må inneholde data
- Datatyper for elementer og attributter
- Standard og faste verdier for elementer og attributter

Prøv selv

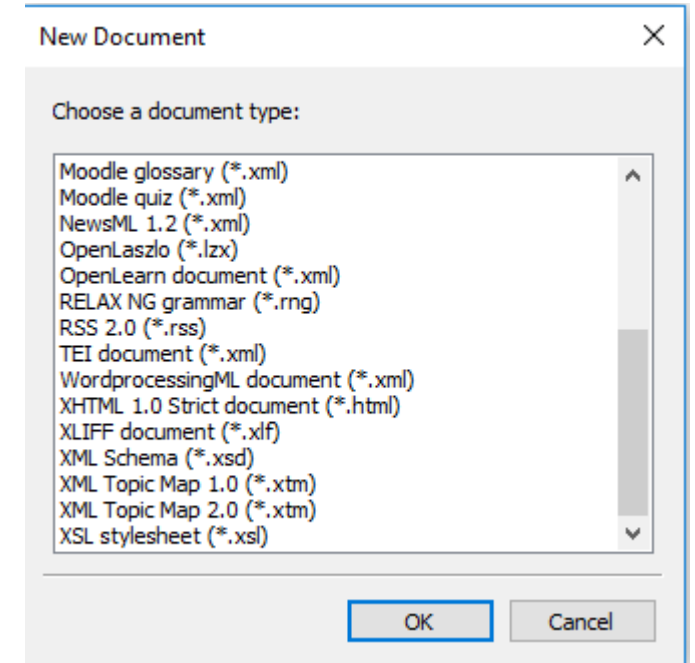
- Åpner XML Copy Editor
- Lag et velformet XML-dokument med opplysningene fra tabellen.
Bruk et attributt til kundenr.

kundenr	123	124
fornavn	Ole	Kari
etternavn	Olsen	Hansen
fødselsdato	17. mai 1970	22. desember 1973
kjønn	Mann	Kvinne

Prøv selv

- Opprett et XML-schema i XML Copy Editor
- File > New > *Velg XML Schema (*.xsd)* fra listen
- Lagrer filen uten å endre, som personer.xsd

```
personer.xml x books2.xml x books2.xsl x books2.xsd x person.xsd x
1 <?xml version="1.0" encoding="UTF-8"?>
2 <xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" elementFormDefault="qualified">
3   <xs:element name="root">
4     <xs:complexType>
5       <xs:sequence>
6         <xs:element name="element" type="xs:string" minOccurs="0" maxOccurs="unbounded"/>
7       </xs:sequence>
8     </xs:complexType>
9   </xs:element>
10 </xs:schema>
11
```



Prøv selv

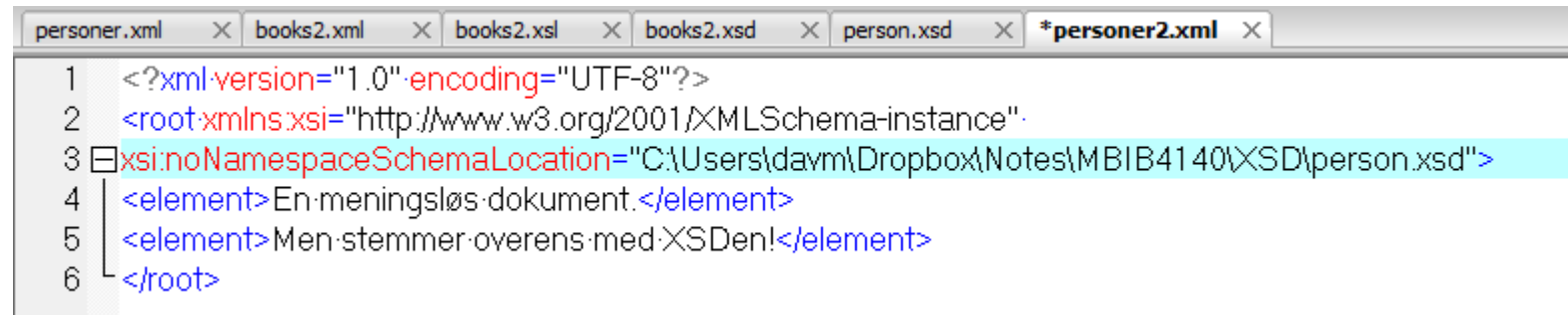
- Opprett en XML dokument med følgende innhold:

```
<?xml version="1.0" encoding="UTF-8"?>  
<root>  
<element>En meningsløs dokument.</element>  
<element>Men stemmer overens med XSDen!</element>  
</root>
```

- Lagre dokumentet som personer2.xml

Prøv selv

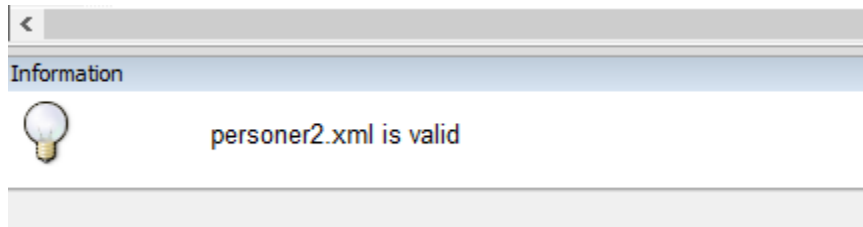
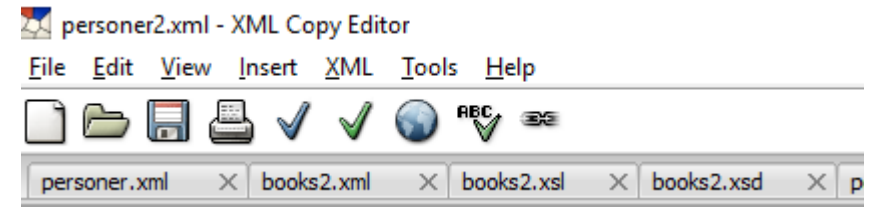
- Koble XSDen til XMLen på følgende måte:
- Ha personer2.xml åpent og velg XML > Associate > XML Schema...
- Finn frem til personer.xsd
- Resultatet skal ligne dette:



```
personer.xml x books2.xml x books2.xsl x books2.xsd x person.xsd x *personer2.xml x
1 <?xml:version="1.0" encoding="UTF-8"?>
2 <root xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
3  xsi:noNamespaceSchemaLocation="C:\Users\daym\Dropbox\Notes\MBIB4140\XSD\person.xsd">
4   <element>En meningsløs dokument.</element>
5   <element>Men stemmer overens med XSDen!</element>
6 </root>
```

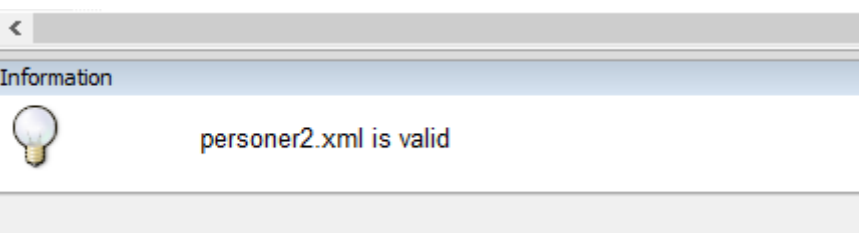
Prøv selv

- Valider personer2.xml
 - XML > Validate > DTD/XML Schema
 - Eller F5
 - Eller ✓
- Resultatet vises under XML-dokumentet:



XML

```
<?xml version="1.0" encoding="UTF-8"?>
<root>
<element>En meningsløs dokument.</element>
<element>Men stemmer overens med
XSDen!</element>
</root>
```



XSD

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema
xmlns:xs="http://www.w3.org/2001/XMLSchema"
elementFormDefault="qualified">
  <xs:element name="root">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="element" type="xs:string"
minOccurs="0" maxOccurs="unbounded"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema
xmlns:xs="http://www.w3.org/2001/XMLSchema"
elementFormDefault="qualified">
  <xs:element name="root">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="element" type="xs:string"
minOccurs="0" maxOccurs="unbounded"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

Elementer:

- complexType
- simpleType

simpleType

- Kan ikke inneholde andre elementer, kun data.
- Kan heller ikke ha attributter.

Hva heter elementet?

Hvilken datatype inneholder elementet?

```
<xs:element name="element" type="xs:string" minOccurs="0" maxOccurs="unbounded"/>
```

Hvor mange ganger kan elementet forekomme (maksimum)?

Hvor mange ganger kan elementet forekomme (minimum)?

simpleType

```
<xs:element name="element" type="xs:string" minOccurs="0"
maxOccurs="unbounded"/>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<root>
</root>
```



```
<?xml version="1.0" encoding="UTF-8"?>
<root>
<element>En</element>
<element>meningsløs</element>
<element>dokument.</element>
<element>Men stemmer overens med
XSDen!</element>
</root>
```



Datatyper

- String (Robert De Niro)
- Integer (1236)
- Decimal (67.2)
- Date (ÅÅÅÅ-MM-DD, f.eks.: 1970-05-17)
- Time (TT:MM:SS, f.eks.: 08:30:00)
- dateTime (ÅÅÅÅ-MM-DDTTT:MM:SS, f.eks.: 1970-05-17T08:30:00)
- Flere datatyper:

https://www.w3schools.com/xml/schema_dtypes_string.asp

XSD Data

XSD String

XSD Date

XSD Numeric

XSD Misc

XSD Reference

Prøv selv

- Endre personer.xsd til dette:

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" elementFormDefault="qualified">
  <xs:element name="person">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="fdato" type="xs:date" minOccurs="0" maxOccurs="1"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

Prøv selv

- Husk å lagre XSDen
- Endre personer2.xml slik at det er gyldig i forhold til den oppdaterte XSDen

complexType

- Elementer som inneholder andre elementer. I dette eksempelet:
person

```
<xs:element name="person">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="fornavn" type="xs:string" />
      <xs:element name="etternavn" type="xs:string" />
      <xs:element name="kjønn" type="xs:string" />
      <xs:element name="fdato" type="xs:date" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
```

Hvis minOccurs eller
maxOccurs mangler brukes 1
som standardverdier.

xs:sequence betyr at underordnede elementer til
person (dvs. fornavn, etternavn og fdato) må
forekomme i denne rekkeløge hvis de finnes.

Prøv selv

- Endre personer.xsd til dette:

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
  elementFormDefault="qualified">
  <xs:element name="person">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="fornavn" type="xs:string" />
        <xs:element name="etternavn" type="xs:string" />
        <xs:element name="kjønn" type="xs:string" />
        <xs:element name="fdato" type="xs:date" />
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

Prøv selv

- Husk å lagre XSDen
- Oppdater personer2.xml slik at den er gyldig i forhold til XSDen
- Hva skjer hvis du endrer rekkefølgen av elementene?
- Hva skjer hvis du legger inn *17. mai 1970* som fødselsdato?
- Hva skjer om du utelater fødselsdato?
- Hvordan endrer du XSDen slik at fødselsdato er valgfritt?

complexType

- xs:all kan brukes istedenfor xs:sequence for å ha valgfri rekkefølge av underordnede elementer.

```
<xs:element name="person">
  <xs:complexType>
    <xs:all>
      <xs:element name="fornavn" type="xs:string" />
      <xs:element name="etternavn" type="xs:string" />
      <xs:element name="kjønn" type="xs:string" />
      <xs:element name="fdato" type="xs:date" />
    </xs:all>
  </xs:complexType>
</xs:element>
```

XSD

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
elementFormDefault="qualified">
  <xs:element name="personer">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="person" type="personInfo"
minOccurs="1" maxOccurs="unbounded" />
      </xs:sequence>
    </xs:complexType>
  </xs:element>

  <xs:complexType name="personInfo">
    <xs:sequence>
      <xs:element name="fornavn" type="xs:string" />
      <xs:element name="etternavn" type="xs:string" />
      <xs:element name="kjønn" type="xs:string" />
      <xs:element name="fdato" type="xs:date" />
    </xs:sequence>
  </xs:complexType>
</xs:schema>
```

XML

```
<personer>
  <person>
    <fornavn></fornavn>
    <etternavn></etternavn>
    <kjønn></kjønn>
    <fdato></fdato>
  </person>
  <person>
    <fornavn></fornavn>
    ...
  </person>
  ...
</personer>
```

Prøv selv

- Oppdater personer.xsd med reglene fra forrige lysark.
- Oppdater personer2.xml med personopplysninger fra dagens første oppgave (glem kundenr foreløpig). Sjekk at XMLen er gyldig i forhold til XSDen.


Attributter

```
<xs:complexType name="personInfo">
  <xs:sequence>
    <xs:element name="fornavn" type="xs:string" />
    <xs:element name="etternavn" type="xs:string" />
    <xs:element name="kjønn" type="xs:string" />
    <xs:element name="fdato" type="xs:date" />
  </xs:sequence>
  <xs:attribute name="kundenr" type="xs:integer" use="required" />
</xs:complexType>
```

Kun for complexType. Legges til etter xs:sequence (evt. xs:all).

```
<personer>
  <person kundenr="123">
    <fornavn></fornavn>
    <etternavn></etternavn>
    <kjønn></kjønn>
    <fdato></fdato>
  </person>
  ...
</personer>
```

Fjern **use="required"** for å gjøre attributtet valgfritt.



Prøv selv

- Oppdater XSDen med reglen for attributtet kundenr
- Legg til kundenr i XMLen og valider

Tomme elementer som kun ha attributter

```
<xs:element name="timezone">  
  <xs:complexType>  
    <xs:attribute name="id" type="xs:string" />  
    <xs:attribute name="utcoffsetMinutes" type="xs:integer" />  
  </xs:complexType>  
</xs:element>
```

```
<timezone id="Europe/Oslo" utcoffsetMinutes="120"/>
```

Restrictions

- **enumeration** kan brukes for å spesifisere en liste av lovlige verdier.

```
<xs:element name="kjønn">  
  <xs:simpleType>  
    <xs:restriction base="xs:string">  
      <xs:enumeration value="Mann" />  
      <xs:enumeration value="Kvinne" />  
    </xs:restriction>  
  </xs:simpleType>  
</xs:element>
```



```
<person>  
  <fornavn>Ola</fornavn>  
  <etternavn>Olsen</etternavn>  
  <kjønn>M</kjønn>  
  <fdato>1970-05-17</fdato>  
</person>
```



```
<person>  
  <fornavn>Ola</fornavn>  
  <etternavn>Olsen</etternavn>  
  <kjønn>Mann</kjønn>  
  <fdato>1970-05-17</fdato>  
</person>
```

Restrictions

```
<xs:complexType name="personInfo">
  <xs:sequence>
    <xs:element name="fornavn" type="xs:string" />
    <xs:element name="etternavn" type="xs:string" />
    <xs:element name="kjønn">
      <xs:simpleType>
        <xs:restriction base="xs:string">
          <xs:enumeration value="Kvinne" />
          <xs:enumeration value="Mann" />
        </xs:restriction>
      </xs:simpleType>
    </xs:element>
    <xs:element name="fdato" type="xs:date" />
  </xs:sequence>
  <xs:attribute name="kundenr" type="xs:integer" use="required" />
</xs:complexType>
```

Prøv selv


- Legg restriction inn i personer.xsd
- Prøv ulike verdier i kjønn og valider

Restrictions


- **length** antall tegn i verdien.

```
<xs:element name="personnummer" minOccurs="1" maxOccurs="1">  
  <xs:simpleType>  
    <xs:restriction base="xs:string">  
      <xs:length value="12" />  
    </xs:restriction>  
  </xs:simpleType>  
</xs:element>
```

```
<person>  
  <fornavn>Ola</fornavn>  
  <etternavn>Olsen</etternavn>  
  <kjønn>Mann</kjønn>  
  <personnummer>170570</personnummer>  
</person>
```



```
<person>  
  <fornavn>Ola</fornavn>  
  <etternavn>Olsen</etternavn>  
  <kjønn>Mann</kjønn>  
  <personnummer>170570171277</personnummer>  
</person>
```




Restrictions

- **pattern** lovlig regex

```
<xs:element name="personnummer" minOccurs="1" maxOccurs="1">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:pattern value="[0-3][0-9][0-1]\d{8}" />
    </xs:restriction>
  </xs:simpleType>
</xs:element>
```

```
<person>
  <fornavn>Ola</fornavn>
  <etternavn>Olsen</etternavn>
  <kjønn>M</kjønn>
  <personnummer>70051771277</personnummer>
</person>
```



Restrictions

- pattern: lovlig regex

```
<xs:element name="kjønn" minOccurs="0" maxOccurs="1">  
  <xs:simpleType>  
    <xs:restriction base="xs:string">  
      <xs:pattern value="Mann|Kvinne" />  
    </xs:restriction>  
  </xs:simpleType>  
</xs:element>
```

Flere regex eksempler:

https://www.w3schools.com/xml/schema_facets.asp

Restrictions

- minLength
- maxLength
- minInclusive
- maxInclusive

Restrictions også for attributtverdier

```
<xs:attribute name="kundenr" use="required">  
  <xs:simpleType>  
    <xs:restriction base="xs:integer">  
      <xs:maxInclusive value="9999" />  
    </xs:restriction>  
  </xs:simpleType>  
</xs:attribute>
```



```
<person kundenr="10000">  
  <fornavn>Ola</fornavn>  
  <etternavn>Olsen</etternavn>  
  <kjønn>M</kjønn>  
</person>
```



```
<person kundenr="12">  
  <fornavn>Ola</fornavn>  
  <etternavn>Olsen</etternavn>  
  <kjønn>M</kjønn>  
</person>
```

Prøv selv

- Last ned: <http://edu.hioa.no/mbib4140/h17/ressurser/xsd/eks-notat.xsd>
- Ta opp XSDen i XML Copy Editor
- Lag et XML-dokument som er gyldig i forhold til XSDen
- Evt. forbedre XSDen
- Eksempelet er fra W3Schools (https://www.w3schools.com/xml/schema_howto.asp)

Prøv selv

- Last ned: <http://edu.hioa.no/mbib4140/h17/ressurser/xsd/eks-bok.xsd>
- Ta opp XSDen i XML Copy Editor
- Lag et XML-dokument som er gyldig i forhold til XSDen
- Evt. forbedre XSDen
- Eksempelet er fra Alexandre Joseph (<https://gist.github.com/jexhson/8011541>)

Prøv selv

- Last ned: <http://edu.hioa.no/mbib4140/h17/ressurser/xml/yr-kikut.xml>
- Ta opp XML-dokumentet i XML Copy Editor
- Lag en XSD som samsvarer med strukturen i XML-dokumentet og valider dokumentet
- Flytt på kommentartegnene (dvs <!--) etterhvert som du utvider XSDen

Neste uke

- XQuery og XPath