

# DAT107 – Databaser, NoSQL

## Øving 1: XML & XML Schema

### Oppgave 1

Finn 6 feil i XML dokumentet under:

```
<xml version="1.0" encoding="UTF-8">
<AUTHOR>Donald Trump</AUTHOR>
<FAQ>
  <INFO>
    <SUBJECT>XML</SUBJECT>
    <VERSION>1.0</VERSION>
    <DATE>20.jun.97<DATE>
    <PART NO="1">
  </INFO>

  <Q NO="1">
    <QTEXT>Hva er XML?</QTEXT>
    <A>Dette lærer vi I DAT107.
  </Q>

  <Q NO="2">
    Et spørsmål til:
    <QTEXT>Hva kan jeg bruke det til?</QTEXT>
    <A>Hva føler du det kan brukes til.</ANSWER>
  </Q>
</PART>
</FAQ>
```

Annotations:

- Missing «?» in <?xml .... ?>
- 2 root elements
- Date format not YYYY-MM-DD
- Not ending element
- Overlapping
- Missing end element
- Mismatch in name of end element.

### Løsning:

Linje 1: Mangler «?» i deklarasjon. <?xml version="1.0" encoding="UTF-8"?>

Linje 2/3: To root element . Kan bare være ETT rotelement.

Linje 7: To «DATE» start elementer, i stedet for at det siste er et slutt element. I tillegg så burde nok dato benyttet formatet YYYY-MM-DD. MEN det er lov å skrive set sånn som det står.

Linje 9/10: Overlappende elementer. «PART» element starter før «INFO» avsluttes.

Linje 14: Mangler slutt tag for <A> elementet.

Linje 20: Start og slutt tag/element må ha samme navn.

Eventuelt (Linje 7: DATE følger ikke xs:date format YYYY-MM-DD)

(Løsningene fortsetter på neste side)

## Oppgave 2

Vi ønsker å benytte XML filer for å overføre informasjon om bøker.

Vi ønsker å overføre følgende data.

- ISBN (13 siffer: Vi forenkler til tall mellom 9 780 000 000 000 – 9 799 999 999 999)
- Tittel (tekst - maksimum 100 tegn)
- Forfattere med fornavn (maksimum 50 tegn) og etternavn (maksimum 50 tegn)
- Format (Tillatte verdier: Pocket, Innbundet, E-bok, Lydbok)
- Utgivelsesår (Tillatte verdier fra 1970 til 2100)
- Forlag (Tekst – maksimum 100 tegn)
- Språk (Norsk, Engelsk)
- Emner (Liste av ord som beskriver boken: f.eks. Java, Programmering)
- Sider (antall sider)

- a) Lag et XML dokument (bøker.xml) med følgende 8 bøker representert ved følgende ISBN nummer fra ARK (<http://www.ark.no>):

**9788269333008, 9788205381056, 9781491915103, 9780321834577,  
9781781101452, 9781491954461, 9788215032511, 9789365890426**

Alle data legges i elementer med unntak av ISBN som legges i et attributt til elementet "bok".

**Løsning: Se «bøker.xml»**

- b) Lag et XML Schema (bøker.xsd) for å validere XML dokumenter av typen beskrevet over.

NB! ISBN er bygd opp etter et spesielt mønster der man tidligere brukte 10 siffer, men i dag benytter 13 siffer. De tre første sifrene i et ISBN-13 er 978 eller 979. Vi skal ikke gjøre våre regler noe mere komplisert enn dette.

**Løsning: Se «bøker.xsd»**

- c) Bruk XML Schema fra oppgave «b» til å validere XML dokumentet fra oppgave «a».

**Løsning: Se attributter til root element i «bøker.xml»**

(Løsningene fortsetter på neste side)

## Oppgave 3

Lag et XML-dokument (oppskrifter.xml) med minimum 3 oppskrifter som er koblet til og benytter XML Schema i filen «oppskrifter.xsd».

**Løsning: Se «oppskrifter.xml»**