Pengantar XML

Yudi Wibisono (<u>yudi@upi.edu</u>) versi dok: Okt 2017

Contoh XML

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<daftar pengarang>
  <pengarang>
     <nama>budi martami</nama>
     <alamat>sarijadi </alamat>
  </pengarang>
  <pengarang>
     <nama>elfan noviari</nama>
     <alamat>gerlong</alamat>
  </pengarang>
</daftar pengarang>
```

Mengapa Belajar XML?

- Standard de-facto untuk bertukar data, menyimpan data, mendeskripsikan data.
- Di-support oleh banyak tools (RDBMS, bahasa pemrograman)
- Webservice → XML pada HTTP
- Platform independen, language independen.
- XML vs JSON?

XML vs JSON

```
"id": 123,
  "title": "Object Thinking",
  "author": "David West",
  "published": {
    "by": "Microsoft Press",
    "year": 2004
}
```

A similar document would look like this in XML (167 characters):

```
<?xml version="1.0"?>
  <book id="123">
        <title>Object Thinking</title>
        <author>David West</author>
        <published>
            <by>Microsoft Press</by>
            year>2004

</published>
        <book>
```

XML vs JSON

- JSON: format data
- XML: Bahasa
 - Attribute & namespace
 - Schema
 - XSL
 - XPATH

Apa XML?

- XML = EXtensible Markup Language
- Markup language untuk dokumen yang mengandung informasi terstruktur.
- Ditujukan untuk pertukaran dokumen via web.
- Berbentuk file teks → Cross platform, indepedenden terhadap software dan hardware.

Aplikasi XML

- Mendeskripsikan dokumen.
- Pertukaran data
- Database

Isi XML

- Tag: <nama> budi </nama>
 - Tag didefinisikan sendiri.
- Tag dapat mempunyai atribut
 - <daftar_mahasiswa jumlah="30"> ... </daftar_mahasiswa>
- Struktur → hirarki (tag dalam tag)

<mahasiswa><nama> budi </nama></mahasiswa>

XML dan HTML

- Persamaan dengan HTML?
 - Sama-sama markup language
- Perbedaan dengan HTML?
 - Semantik HTML telah didefinisikan. Contoh:

 untuk bold
 - XML lebih ketat aturannya. Contoh: setiap tag harus ada penutupnya (<tag> </tag>)
 - XML case sensitive

Keuntungan XML

- Self Documenting → dengan melihat tag, dapat diketahui isi dokumen.
- Dapat dibaca software dan manusia.
- Fleksibel
- Dapat dikembangkan tanpa melanggar format lama.
 - Contoh: pada <pengarang>, tambah tag
 <telepon>, maka XML yang lama tetap dapat dibaca.

Keuntungan XML (lanj)

- Hirarkis → dapat merepresentasikan data kompleks
- Independen terhadap bahasa pemrograman, OS

Kerugian XML

 Pengulangan tag → tidak efisien, ukuran membengkak

Beberapa Aturan XML

- Setiap tag harus ada penutupnya.
- Penamaan Tag
 - case sensitive. <nama> tidak sama dengan <Nama>.
 - Tidak diawali dengan angka
 - Tidak mengandung spasi
 - Hindari '-' dan '.'
- Urutan hirarki harus benar.

Contoh yang salah:

<mahasiswa> <nama> yudi </mahasiswa></nama>

Aturan XML (lanj)

Setiap XML harus mengandung root (akar)

Setiap atribut harus dalam tanda petik. Contoh:

```
<surat tanggal = "12/12/2007">
<catatan oleh = " Jum'at " >
<catatan oleh = ' Budi "si cepat" Martami '>
```

Komentar dalam XML:

```
<!-- ini komentar lho --!>
```

Elemen: tag, isi, sampai tag penutup.
 <tag> isi tag </tag>

Elemen dan Atribut

atribut vs elemen, mana yang lebih baik?

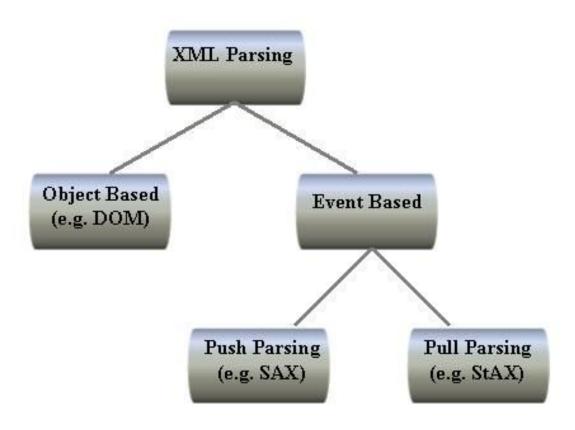
ATAU

```
<surat>
<tanggal> 2/2/2005 </tanggal>
<pengirim> Yudi </pengirim>
</surat>
```

XML Parser

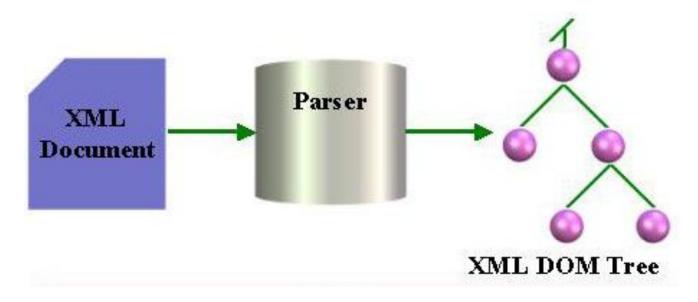
- Parser → program untuk "membaca" dokumen XML
- Dua cara:
 - Tree based, isi XML dipindahkan ke memori terlebih dulu → DOM (Document Object Model).
 - Event based → tidak dipindahkan ke memori.
 Fungsi callback dipanggil setiap menemui
 element. → SAX (Simple API XML)

XML Parser



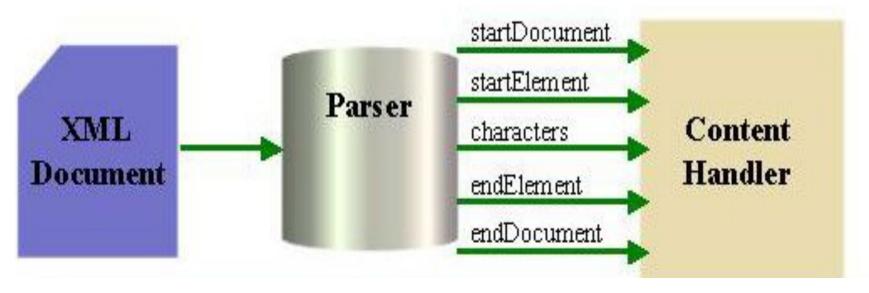
http://www.xml.com/2005/07/06/graphics/image001.jpg

DOM



http://www.xml.com/2005/07/06/graphics/image003.jpg

SAX: Event Based



http://www.xml.com/2005/07/06/graphics/image002.jpg

Spesifikasi

- DTD → Document Type Declaration
 - Definsi Tag dan atribut
- XML Scheme → lebih general dari DTD
- XSL → eXtensible Style Languange
 Bagaimana XML ditampilkan

DTD: Document Type Declaration

- Dari 1970-an
- Well Formed XML → XML yang mengikuti aturan:
 - Mempunyai root
 - Setiap element mempunyai penutup
 - Case sensitive
 - Nilai atribut dalam tanda petik (" ")
 - Tersusun dengan benar (nested properly)

Wellformed belum tentu VALID!

Agar dapat dikatakan VALID harus mengikuti
DTD!

DTD (2)

```
<?xml version="1.0"?>
<!DOCTYPE note [
   <!ELEMENT note (to,from,heading,body)>
   <!ELEMENT to (#PCDATA)>
   <!ELEMENT from (#PCDATA)>
   <!ELEMENT heading (#PCDATA)>
   <!ELEMENT body (#PCDATA)> ]>
<note>
   <to>Akhmad</to>
   <from>Budi</from>
   <heading>UTS</heading>
   <body>Jangan Lupa UTS!</body>
</note>
```

DTD mendefinisikan elemen yang legal dalam sebuah dokumen XML → mengapa menggunakan DTD?

DTD Construct

- ELEMENT
- ATTLIST
- ENTITY
- NOTATION

DTD (2)

DTD Sebagai file external:

DTD: Element

```
<!ELEMENT Keranjang ( Jeruk, (Apel | Jambu)) >
Maka contoh XML-nya yang valid :
<Keranjang>
   <Jeruk> </Jeruk>
   <Apel></Apel>
</Keranjang>
<Keranjang>
   <Jeruk></Jeruk>
   <Jambu></Jambu>
</Keranjang>
```

DTD: Element (kardinalitas)

```
<!ELEMENT Keranjang ( Jeruk+, (Apel | Jambu)* ) >
Contoh XML yang valid:
                                 ?: Opsional, bisa ada atau tidak *
<Keranjang>
                                  : Nol atau banyak
   <Jeruk> </Jeruk>
                                  + : Satu atau banyak
</Keranjang>
<Keranjang>
    <Jeruk></Jeruk>
   <Jambu></Jambu>
    <Apel></Apel>
    <Apel></Apel>
</Keranjang>
```

DTD: ATRIBUT

Contoh 1:

```
<!ATTLIST Keranjang warna CDATA "hitam" #IMPLIED> XML-nya:
```

```
<Keranjang warna = "biru"> ... </Keranjang>
atau
```

<Keranjang> </Keranjang>

Implied artinya tidak wajib, jika kosong menggunakan nilai default yaitu "hitam"

Contoh 2:

```
<!ATTLIST Buku ISBN ID #REQUIRED>
```

(ID artinya ISBN harus unik dan REQUIRED artinya atribut ini wajib ada)

DTD: Entitas

Shortcut dari teks yang umum (konstanta)

```
<!ENTITY bm "Budi Martami>
```

di XML-nya gunakan &entity;

```
<Pengarang> &bm; </Pengarang>
```

DTD:Notation

Untuk data selain XML

```
<!NOTATION jpg SYSTEM "jpgviewer.exe"> <!ATTLIST Image type NOTATION (gif|jpg)>
```

XML:

```
<lmage type ="jpg"> .... </image>
```

Newspaper Article DTD

Copied from http://www.vervet.com/

```
<!DOCTYPE NEWSPAPER [</pre>
<!ELEMENT NEWSPAPER (ARTICLE+)>
<!ELEMENT ARTICLE (HEADLINE,BYLINE,LEAD,BODY,NOTES)>
<!ELEMENT HEADLINE (#PCDATA)>
<!ELEMENT BYLINE (#PCDATA)>
<!ELEMENT LEAD (#PCDATA)>
<!ELEMENT BODY (#PCDATA)>
<!ELEMENT NOTES (#PCDATA)>
<!ATTLIST ARTICLE AUTHOR CDATA #REOUIRED>
<!ATTLIST ARTICLE EDITOR CDATA #IMPLIED>
<!ATTLIST ARTICLE DATE CDATA #IMPLIED>
<!ATTLIST ARTICLE EDITION CDATA #IMPLIED>
<!ENTITY NEWSPAPER "Vervet Logic Times">
<!ENTITY PUBLISHER "Vervet Logic Press">
K!ENTITY COPYRIGHT "Copyright 1998 Vervet Logic Press">
```

```
<!DOCTYPE CATALOG [
<!ENTITY AUTHOR "John Doe">
<!ENTITY COMPANY "JD Power Tools, Inc.">
<!ENTITY EMAIL "jd@jd-tools.com">
<!ELEMENT CATALOG (PRODUCT+)>
<!ELEMENT PRODUCT
(SPECIFICATIONS+,OPTIONS?,PRICE+,NOTES?)>
<!ATTLIST PRODUCT
NAME CDATA #IMPLIED
CATEGORY (HandTool|Table|Shop-Professional) "HandTool"
PARTNUM CDATA #IMPLIED
PLANT (Pittsburgh|Milwaukee|Chicago) "Chicago"
INVENTORY (InStock|Backordered|Discontinued) "InStock">
<!ELEMENT SPECIFICATIONS (#PCDATA)>
<!ATTLIST SPECIFICATIONS
WEIGHT CDATA #IMPLIED
POWER CDATA #IMPLIED>
<!ELEMENT OPTIONS (#PCDATA)>
<!ATTLIST OPTIONS
FINISH (Metal|Polished|Matte) "Matte"
ADAPTER (Included | Optional | NotApplicable) "Included"
CASE (HardShell|Soft|NotApplicable) "HardShell">
<!ELEMENT PRICE (#PCDATA)>
<!ATTLIST PRICE
MSRP CDATA #IMPLIED
WHOLESALE CDATA #IMPLIED
STREET CDATA #IMPLIED
SHIPPING CDATA #IMPLIED>
<!ELEMENT NOTES (#PCDATA)>
```

Contoh DTD (lanjutan) Katalog Produk

Latihan membuat DTD

- Buatlah DTD XML untuk menyimpan data nilai mahasiswa
- Buatlah DTD XML untuk menyimpan data fasilitas pada suatu hotel

Kelemahan DTD?

- Sintaksnya bukan format XML standard
- Kurang lengkap
- Tidak fleksibel
- Setiap orang dapat membuat tag sendiri yang mungkin bentrok
- Reuse DTD? sulit.
- Tidak mensupport strong typing (integer, string date)

Solusi terbaru: XML Schema

XML Schema

- Dari tahun 2001
- Menggunakan format XML
- Calon kuat pengganti DTD di masa depan
- Dikenal juga dengan nama
 - XSD: XML Schema Document
 - WXS: W3C XML Schema
- Organisasi lebih baik (namespace)
- Strong typing → 44 tipe (DTD hanya 10)
- Lebih presisi dan fleksibel (Custom datatype)
 - Contoh: dapat mendefinisikan tipe umur : 0<umur<120
 - Dapat mendefinisikan urutan child
- Compatible dengan teknologi XML yang lain: Web Service, Xquery

XML Schema

Cek schema online:

http://www.utilities-online.info/xsdvalidation

Perbandingan DTD vs XML Schema

(dari http://www.adp-gmbh.ch)

```
<?xml version="1.0" encoding="UTF-8"?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <xsd:element name="uuml" type="xsd:token" fixed="ü"/>
                                                                    XMI Schema
  <xsd:complexType name="name">
    <xsd:sequence>
      <xsd:element name="title"</pre>
                                       type="xsd:string" min0ccurs="0" max0ccurs="1" />
                                       type="xsd:string" minOccurs="1" maxOccurs="1" />
      <xsd:element name="first name"</pre>
      <xsd:element name="middle name" type="xsd:string" min0ccurs="0" max0ccurs="1" />
      <xsd:element name="last name"</pre>
                                       type="xsd:string" min0ccurs="1" max0ccurs="1" />
      <xsd:element name="suffix"</pre>
                                       type="xsd:string" minOccurs="0" maxOccurs="1" />
    </xsd:sequence>
  </xsd:complexType>
</xsd:schema>
```

Struktur Schema

- Pembukaan (Preamble)
- Definisi tipe (simple, kompleks) dan deklarasi elemen
- Content model??
- Model Group??

Namespace

- Format: xmlns:XML_name_space= URI.

```
xmlns:buku = " http://www.upi.edu/buku.xsd "
xmlns:drama = "http://www.upi.edu/drama.xsd "
```

```
<buku:pengarang> ... </buku:pengarang>
<drama:pengarang> ... </drama:pengarang>
```

Definisi Type

- type ="[jenis tipe]"
 Jenis tipe: date, string, integer, double
- <tgl type = "date"> 2004-03-11 <tgl>
- Format tipe "date" adalah YYYY-MM-DD

Definisi Custom Tipe & Elemen: Simple

Tipe umur dimana 0 < umur <= 120:

```
<xsd:simpleType name="tipeUmur">
    <xsd:restriction base="xsd:positiveInteger">
        <xsd:minExclusive value="0"/>
        <xsd:maxInclusive value="120"/>
    </xsd:restriction>
</xsd:simpleType>
<xsd:element name="umur" type="tipeUmur"/>
Tanggal di bulan Mei YYYY – 05 - DD:
<xsd:simpleType name="may date">
      <xsd:restriction base="date">
            < xsd: pattern value = "\d{4}-05-\d{2}"/>
      </xsd:restriction>
</xsd:simpleType>
```

Pembukaan

```
<?xml version="1.0"?>
< xsd:schema
 xmlns:xsd="http://www.w3.org/2001/XM
 LSchema ">
</xsd:schema>
Root: "schema".
```

xsd adalah namespace

Deklarasi Element

```
<element name="kodepos" type = "string" />
<element name="kecepatan" type = "float" />
Tipe yang lain:
```

- "int"
- · "boolean"
- "double"
- "decimal"
- "duration"
- "dateTime"
- "time"
- "date"

Contoh Definisi Tipe, Elemen, Atribut: Tipe Kompleks

XML:

```
<note ID="5">
     <to>Akhmad</to>
     <date>2005-12-30</date>
</note>
```

DTD:

```
<!ELEMENT note (to,date)>
<!ELEMENT to (#PCDATA)>
<!ELEMENT date (#PCDATA)>
<!ATTLIST note ID #REQUIRED>
```

Definisi Tipe, Elemen, Atribut: Tipe Kompleks (Lanj)

XML Schema:

Definisi Tipe : Tipe Kompleks (lanj)

```
DTD:
<!ELEMENT Keranjang ( Jeruk+, Apel*) >
XML Scheme:
<xsd:element name = "Keranjang">
  <xsd:complexType>
     <xsd:sequence>
      <xsd:element name = "Jeruk" type "xsd:string" minOccurs = "1"</pre>
  maxOccurs="unbounded"/>
       <xsd:element name = "Apel" type "xsd:string" minOccurs = "0"</pre>
  maxOccurs="unbounded"/>
     </xsd:sequence>
  </xsd:complexType>
</xsd:element>
```

Definisi Tipe : Tipe Kompleks (choice)

```
DTD:
<!ELEMENT Keranjang (Jeruk|Apel) >
XML Scheme:
<xsd:element name = "Keranjang">
  <xsd:complexType>
    <xsd:choice>
      <xsd:element name = "Jeruk" type ="xsd:string"/>
      <xsd:element name = "Apel" type = "xsd:string"/>
    </xsd:choice>
  </xsd:complexType>
</xsd:element>
```

Definisi Tipe: Tipe Kompleks (group

```
<!ELEMENT Keranjang ( Jeruk+, (Apel | Jambu)* )>
XML Scheme:
<xsd:group name="apeljambu">
    <xsd:choice>
      <xsd:element name="Apel" type="xsd:string"/>
      <xsd:element name="Jambu" type="xsd:string"/>
   </xsd:choice>
</xsd:group>
<xsd:element name = "Keranjang">
  <xsd:complexType>
     <xsd: sequence>
       <xsd:element name = "Jeruk" type = "xsd:string" minOccurs = "1"</pre>
   maxOccurs="unbounded />
       <xsd:group ref = "apeljambu" minOccurs = "0" maxOccurs="unbounded"/>
     </xsd:sequence>
  </xsd:complexType>
</xsd:element>
```

Penggunaan Ref

```
<xsd:element name = "note">
    <xsd:complexType>
        <xsd:sequence>
          <xsd:element name = "to" type = "xsd:string"/>
      <xsd:element name = "date" type = "xsd:date"/>
        </xsd:sequence>
    </xsd:complexType>
</xsd:element>
dapat ditulis:
<xsd:element name = "to" type = "xsd:string"/>
<xsd:element name = "date" type = "xsd:date"/>
<xsd:element name = "note">
    <xsd:complexType>
        <xsd:sequence>
          <xsd:element ref = "to"/>
          <xsd:element ref = "date"/>
        </xsd:sequence>
    </xsd:complexType>
</xsd:element>
```

Contoh XML dan Schema-nya

(http://www.xml.com/pub/a/2000/11/29/schemas/part1.html?page=1)

XML

```
<?xml version="1.0" encoding="UTF-8"?>
<book isbn="0836217462">
<title>
 Being a Dog Is a Full-Time Job
</title>
 <author>Charles M. Schulz</author>
 <character>
  <name>Snoopy</name>
  <friend-of>Peppermint Patty</friend-of>
  <since>1950-10-04</since>
  <qualification>
   extroverted beagle
  </gualification>
 </character>
 <character>
  <name>Peppermint Pattv</name>
  <since>1966-08-22</since>
  <qualification>bold, brash and tomboyish</qua
 </character>
</book>
```

```
Schema
<?xml version="1.0" encoding="utf-8"?>
xs:schema xmlns:xs="http://www.w3.org/2001/XM
  <xs:element name="book">
    <xs:complexType>
      <xs:sequence>
         <xs:element name="title" type="xs:stri</pre>
         <xs:element name="author" type="xs:str</pre>
         <xs:element name="character" minOccurs</pre>
           <xs:complexTvpe>
             <xs:sequence>
               <xs:element name="name" type="xs</pre>
               <xs:element name="friend-of" typ</pre>
                                  maxOccurs="unbo
               <xs:element name="since" type="x</pre>
               <xs:element name="qualification"</pre>
             </xs:sequence>
           </xs:complexType>
         </xs:element>
      </xs:sequence>
      <xs:attribute name="isbn" type="xs:strin</pre>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

```
<?xml version="1.0" encoding="utf-8"?>
xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="book">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="title" type="xs:string"/>
        <xs:element name="author" type="xs:string"/>
        <xs:element name="character" minOccurs="0" maxOccurs="unbounded">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="name" type="xs:string"/>
              <xs:element name="friend-of" type="xs:string" minOccurs="0"</pre>
                                maxOccurs="unbounded"/>
              <xs:element name="since" type="xs:date"/>
              <xs:element name="qualification" type="xs:string"/>
            </xs:sequence>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
      <xs:attribute name="isbn" type="xs:string"/>
    </xs:complexType>
  </xs:element>
                          yual@upi.eau
                                                        UC
```

Contoh XML dan XML-Schema(dari http://lucas.ucs.ed.ac.uk/xml-schema/)

```
a sample instance document: file landrover.xml
<?xml version = "1.0" encoding = "UTF-8"?>
vehicles
 xmlns:xsi = "http://www.w3.org/2001/XMLSchema-instance"
 xsi:noNamespaceSchemaLocation = "http://lucas.ucs.ed.ac.uk/xml-schema/xmlns/landrover.xsd">
                                      The Landrover schema: file landrover.xsd
<vehicle>
                                                                                               Schema hal 1
 <nickname>Count Zero</nickname>
                                      <?xml version = "1.0" encoding = "UTF-8"?>
 <model>Series I, 80"</model>
                                      <xsd:schema xmlns:xsd = "http://www.w3.org/2001/XMLSchema">
                                                                                                           <xsd:element name = "start">
 <construction>
                                                                                                                                                hal 2
                                                                                                            <xsd:complexType>
  <start>
                                                                                                            <xsd:sequence>
                                        <xsd:element name = "vehicles">
   <dom>21</dom>
                                                                                                             <xsd:element ref = "dom"/>
                                         <xsd:complexType>
   <month>July</month>
                                                                                                              <xsd:element ref = "month"/>
                                          <xsd:sequence>
   <year>1949</year>
                                                                                                             <xsd:element ref = "year"/>
                                           <xsd:element ref = "vehicle" maxOccurs = "unbounded"/</pre>
  </start>
                                                                                                            </xsd:sequence>
                                          </xsd:sequence>
  <end>
                                                                                                            </xsd:complexTvpe>
                                         </xsd:complexType>
                                                                                                           </xsd:element>
   <dom>9</dom>
                                        </xsd:element>
   <month>August</month>
                                                                                                           <xsd:element name = "end">
   <year>1949
                                        <xsd:element name = "vehicle">
                                                                                                            <xsd:complexType>
  </end>
                                         <xsd:complexType>
                                                                                                            <xsd:sequence>
 </construction>
                                          <xsd:sequence>
                                                                                                             <xsd:element ref = "dom"/>
                                           <xsd:element name = "nickname" type = "xsd:string" xaxOccurs</pre>
 <mods>
                                                                                                             <xsd:element ref = "month"/>
  <mod>Change Engine</mod>
                                           <xsd:element name = "model" type = "xsd:string" \( \begin{align*} \begin{align*} \text{ } \\ \text{ } \end{align*} \]</pre>
                                                                                                             <xsd:element ref = "vear"/>
  <mod>Change pedals</mod>
                                           <xsd:element name = "construction">
                                                                                                            </xsd:sequence>
  <mod>Change gearbox</mod>
                                            <xsd:complexType>
                                                                                                            </xsd:complexType>
  <mod>Fit Rollcage</mod>
                                             <xsd:sequence>
                                                                                                           </xsd:element>
                                              <xsd:element ref = "start"/>
 </mods>
                                              <xsd:element ref = "end"/>
</re>
                                                                                                           <xsd:element name = "dom" type = "xsd:string"/>
                                             </xsd:sequence>
:/vehicles>
                                                                                                           <xsd:element name = "month" type = "xsd:string"/>
                                            </xsd:complexType>
                                                                                                           <xsd:element name = "year" type = "xsd:string"/>
                                           </xsd:element>
                                           <xsd:element name = "mods">
                                                                                                          xsd:schema>
                                            <xsd:complexType>
                                             <xsd:sequence>
                                              <xsd:element name = "mod" type = "xsd:string" maxOccurs = "unbounded"/>
                                             </xsd:sequence>
                                            </xsd:complexType>
                                           </xsd:element>
                                          </xsd:sequence>
                                                                                                   Bersambung ke hal 2
```

09/03/2017

</xsd:complexType> </xsd:element>

```
<?xml version="1.0" encoding="UTF-8"?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <xsd:element name="uuml" type="xsd:token" fixed="ü"/>
  <xsd:complexType name="name">
    <xsd:sequence>
      <xsd:element name="title"</pre>
                                        type="xsd:string" min0ccurs="0" max0ccurs="1" />
      <xsd:element name="first name"</pre>
                                        type="xsd:string" min0ccurs="1" max0ccurs="1" />
      <xsd:element name="middle name"</pre>
                                        type="xsd:string" min0ccurs="0" max0ccurs="1" />
      <xsd:element name="last name"</pre>
                                        type="xsd:string" min0ccurs="1" max0ccurs="1" />
      <xsd:element name="suffix"</pre>
                                        type="xsd:string" min0ccurs="0" max0ccurs="1" />
    </xsd:sequence>
  </xsd:complexType>
</xsd:schema>
```

todo

DTD dibuang... check schema online: