

# XML CATALOGS

Developer Meetup, February 23rd

WHAT IS IT AND WHY SHOULD I BE USING IT?

# XML SCHEMAS

```
<?xml version="1.0" encoding="UTF-8"?>
<Dokument id="idd38838df-7255-4705-ba5f-7b644c48dc6c"
  SchemaLocation="http://www.retsinformation.dk/offentlig/xml/schemas/2015/06/2
  <Meta id="idc723f9f4-ddad-4364-8bb6-edef72eece0b">
    <DocumentType>LOV A#LOKDOK02</DocumentType>
    <Rank>A</Rank>
    <AccessionNumber>A20160007930</AccessionNumber>
    <DocumentId>CK000529</DocumentId>
    <UniqueDocumentId>177192</UniqueDocumentId>
    <DocumentTitle>Lov om ophævelse af lov om randzoner</DocumentTitle>
    <Year>2016</Year>
    <DiesSigni>2016-01-27</DiesSigni>
    <DateOfSubmit id="submit_1"/>
    <EndDate REFid="submit_1">2017-01-06</EndDate>
    <Status>Valid</Status>
    <Number>79</Number>
    <AendresAf>
      <Dokument>
        ...
      </Dokument>
    </Dokument>
```

So where's [LOVC.Retsinfo.LexDania\\_2.1.xsd](#)

DTDs

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE LOV PUBLIC "-//Thomson Information AS//DTD LOV//DK"
"allov.dtd">
<LOV DOKVERSTYPE="V" ID="LBKG20131073" IKRAFTDATO="20110407">
  <L>
    <TITEL>Lovbekendtg?relse <DATO>2013-09-04</DATO> <LOVNR>nr. 1073</LOVNR> <N
    ...
  </LOV>
```

So where's **allov.dtd**?

# XSLTs

```
<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet version="1.0"
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
  xmlns:fo="http://www.w3.org/1999/XSL/Format">

  <xsl:import href="http://ns.karnovgroup.com/stylesheets/markdown/kgbook/fo/kg
  ...
</xsl:stylesheet>
```

So where's [ns.karnovgroup.com/.../kgbook.xsl](http://ns.karnovgroup.com/.../kgbook.xsl)?

# CATALOGS!!!

That's where catalogs comes in

```
~/src/xml-toolbox/  
├─ schemas  
|   ├─ catalog-ncr.xml  
|   ├─ catalog-utf-8.xml  
|   ├─ catalog.xml  
|   ├─ fid.xml  
|   ├─ ft.xml  
|   ├─ general.xml  
|   ├─ kgentities-ncr.xml  
|   ├─ kgentities-utf-8.xml  
|   ├─ ns.karnovgroup.com  
|   ├─ oxygenxml  
|   ├─ www.retsinformation.dk  
|   └─ xhtml.xml  
├─ stylesheets  
|   ├─ docbook.sourceforge.net  
|   ├─ ns.karnovgroup.com  
|   └─ www.xsltfunctions.com
```

# CATALOG.XML

```
<?xml version="1.0" encoding="UTF-8"?>
<catalog xmlns="urn:oasis:names:tc:entity:xmlns:xml:catalog">
  <rewriteURI uriStartString="http://ns.karnovgroup.com/stylesheets/"
    rewritePrefix="."/>
  <rewriteSystem systemIdStartString="http://ns.karnovgroup.com/stylesheets/"
    rewritePrefix="."/>
</catalog>
```

NOKOGIRI???

It's just LibXML

Run your ruby stuff with environment variable set

```
XML_CATALOG_FILES='MY-PATH-TO-MY-CATALOG.xml' ruby ...
```

... or use the xml-toolbox



# XML-TOOLBOX

```
require 'nokogiri'  
require 'xml_toolbox'  
XmlToolbox.utf8_catalog_resolver
```

# DEMOS

Play around with `xmlcatalog` (part of the `libxml` package):

```
xmlcatalog path-to-my-catalog.xml public-or-system-identifier
```

This should fail because it cannot find the DTD file

```
ruby nokogiri-without-toolbox.rb VL2000_B_1301_98.xml
```

This should succeed because the catalog resolved it

```
ruby nokogiri-with-toolbox.rb VL2000_B_1301_98.xml
```

This should show DTD errors

```
ruby nokogiri-with-toolbox.rb VL2000_B_1301_98-wrong.xml
```

# DONE!

<http://xmlsoft.org/catalog.html>

<https://github.com/karnov/xml-toolbox/>