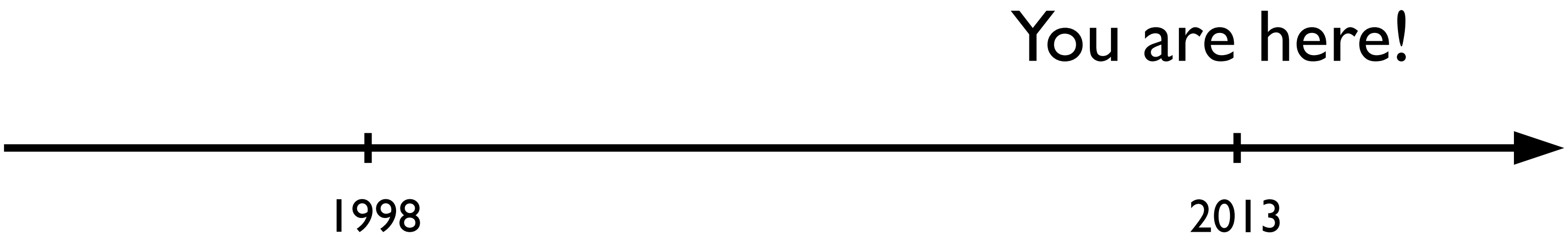
Quo vadis XML?

Maik Stührenberg
Bielefeld University
maik.stuehrenberg@uni-bielefeld.de
@mstuehrenberg

> “Do we need new syntax, data model or both or is XML just fine after 15 years?”
>
> XMLPrague CFP

Celebrating XML's birthday

XML is now 15 years old...



You are here!

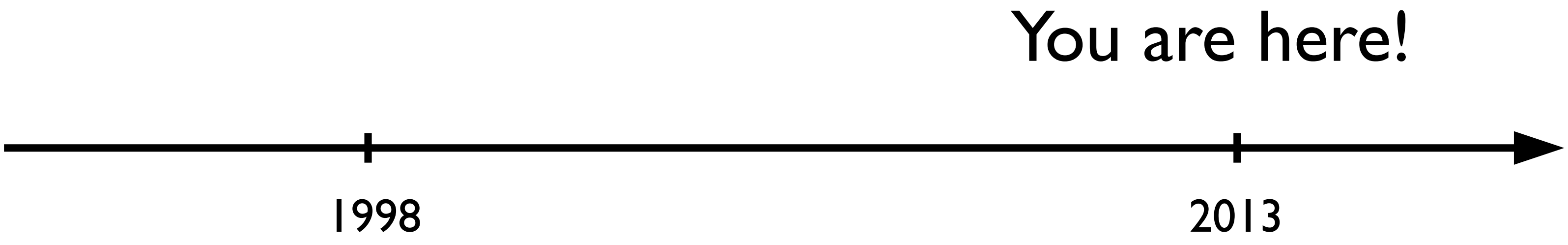
1998

2013

Celebrating XML's birthday

During this time...

- ...XSLT and XPath were invented and extended and modified up to version 3.0
- ...XQuery was developed
- ...XML Schema 1.0 and 1.1 were standardized
- ...RELAX NG was built on the shoulder of RELAX and TREX
- Schematron was standardized



You are here!

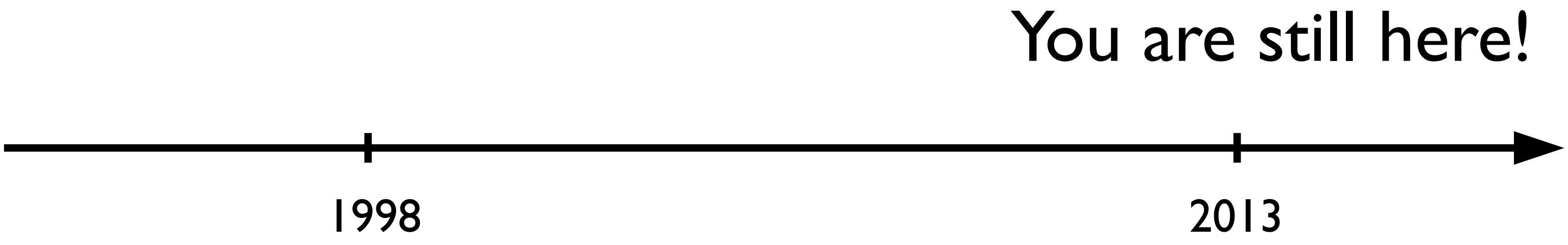
1998

2013

Celebrating XML's birthday

But XML itself...

- ...maintained more or less stable – we have five editions of XML 1.0 and two editions of XML 1.1 (which was not a huge success)



You are still here!

1998

2013

Why is XML so stable? So mature?

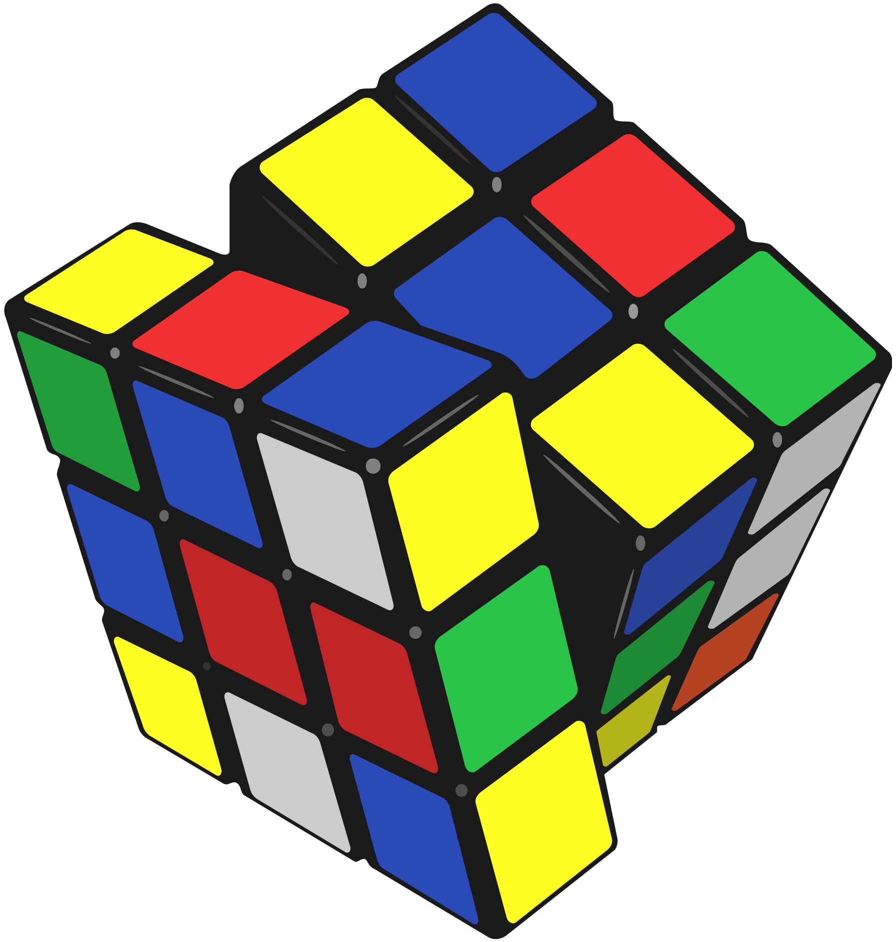
Maybe the answer lies in its history...

XML's heritage

- First considerations about generic markup have been made in the 60s such as GenCode, Text Description Language (TDL, later GML)
- In the 1970s GenCode and GML combined forces

The 80s: SGML

- In the 80s, SGML was standardized as ISO standard
- SGML as a meta language laid the foundation for a whole family of markup languages such as
  - TEI
  - Docbook
  - HTML up to 4.01



The 90s: Lots of Pulp – and HTML and XML

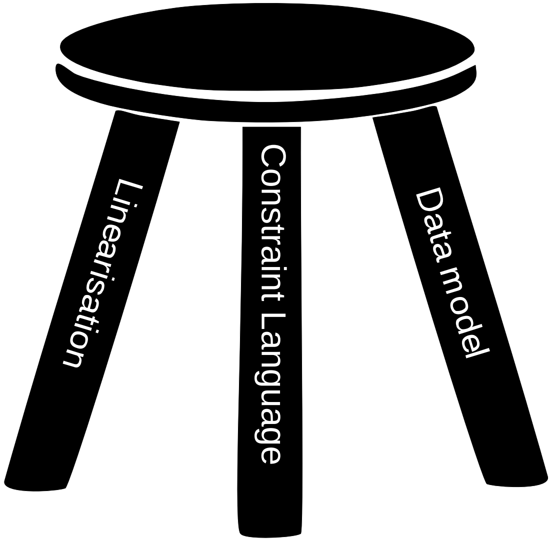
- HTML was developed in the early 90s based on SGML
- XML was created as an SGML subset in 1997 and standardized as W3C Recommendation in 1998 to conquer the web
- As an SGML subset it inherited most basic components

Basic concepts of a markup language

In 1999 Sperberg-McQueen and Huitfeldt named three basic concepts of SGML

- Linearization
- Data model
- Constraint language

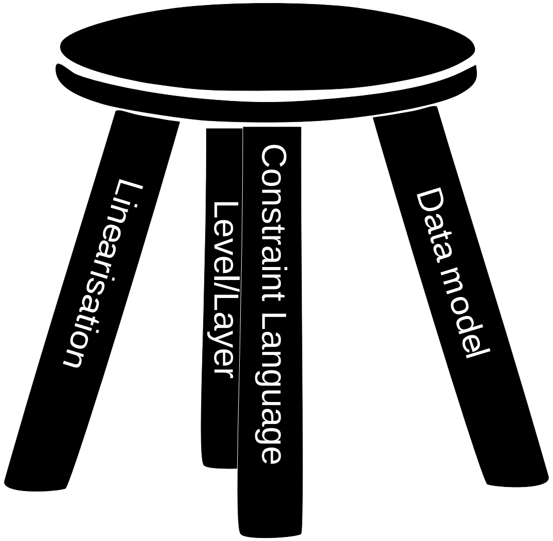
In 2004 the authors coined the tripod metaphor



Basic concepts of a markup language

Witt added 2004 the distinction between

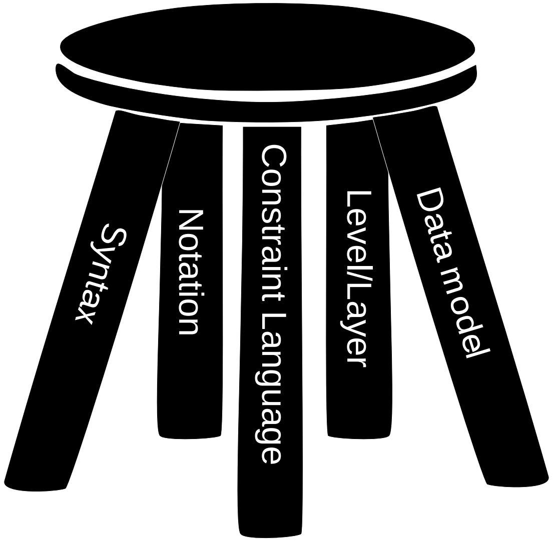
- a level (the concept that drives certain markup) and
- the layer (the serialization of the markup)
  → Think of microformats or TEI feature structures
- This distinction allows for generic
  markup – and a quatropod metaphor



Basic concepts of a markup language

But we can go even further:

- We can split up linearisation into
  - Notation (inline vs standoff) and
  - Syntax (this is especially interesting when dealing with non-XML-based markup language)



In terms of XML

- XML's data model is tree-like → let's stick with the term hierarchy
- XML's syntax consists of elements, attributes and tags (amongst others) and uses special characters to differentiate between markup and content
- XML's notation can be inline or standoff
- Various constraint languages are available
- Level/layer abstraction is grammar-driven

XML is 15 – that is, it's a teenager

Teenage problems:

- Teenagers test boundaries – e. g. boundaries of the data model, such as overlapping markup
- Teenagers do not behave as expected – take the different requirements regarding deterministic content models in DTD, XSD and RNG
- Teenagers challenge traditional values – such as draconian error handling in parsing (see HTML5)
- Teenagers try to mimic others – such as JSON's data model (wait for Eric van der Vlist's talk)

How to cope with teenage problems

Try to educate them

- Find a way to cope with its limitations
  → either try to simplify or extend XML's components without losing compatibility

Start over again

- Tell the teen to get a life – and invent alternative approaches which address at least one of XML's limitations

Approaches that are compatible to teen XML

- MicroXML
- XStandoff
- xLMNL
- ...

Approaches that are incompatible to teen XML

- LMNL
- FtanML
- ...

MicroXML

- We won't go into details here since we've already heard Uche Ogbuji's talk
- MicroXML addresses the following of XML's components:
  - Data model
  - Syntax
  - Parsing

XStandoff

- A hybrid meta markup language that was developed since 2008 to allow multiple (and possible overlapping) annotations – especially linguistically annotated data
- Support for discontinuous elements, multiple parenthood and differentiation between dominance and containment
- XStandoff addresses the following of XML's components:
  - Data model (graph- or GODDAG-like)
  - Level/layer distinction

XStandoff – A simple example

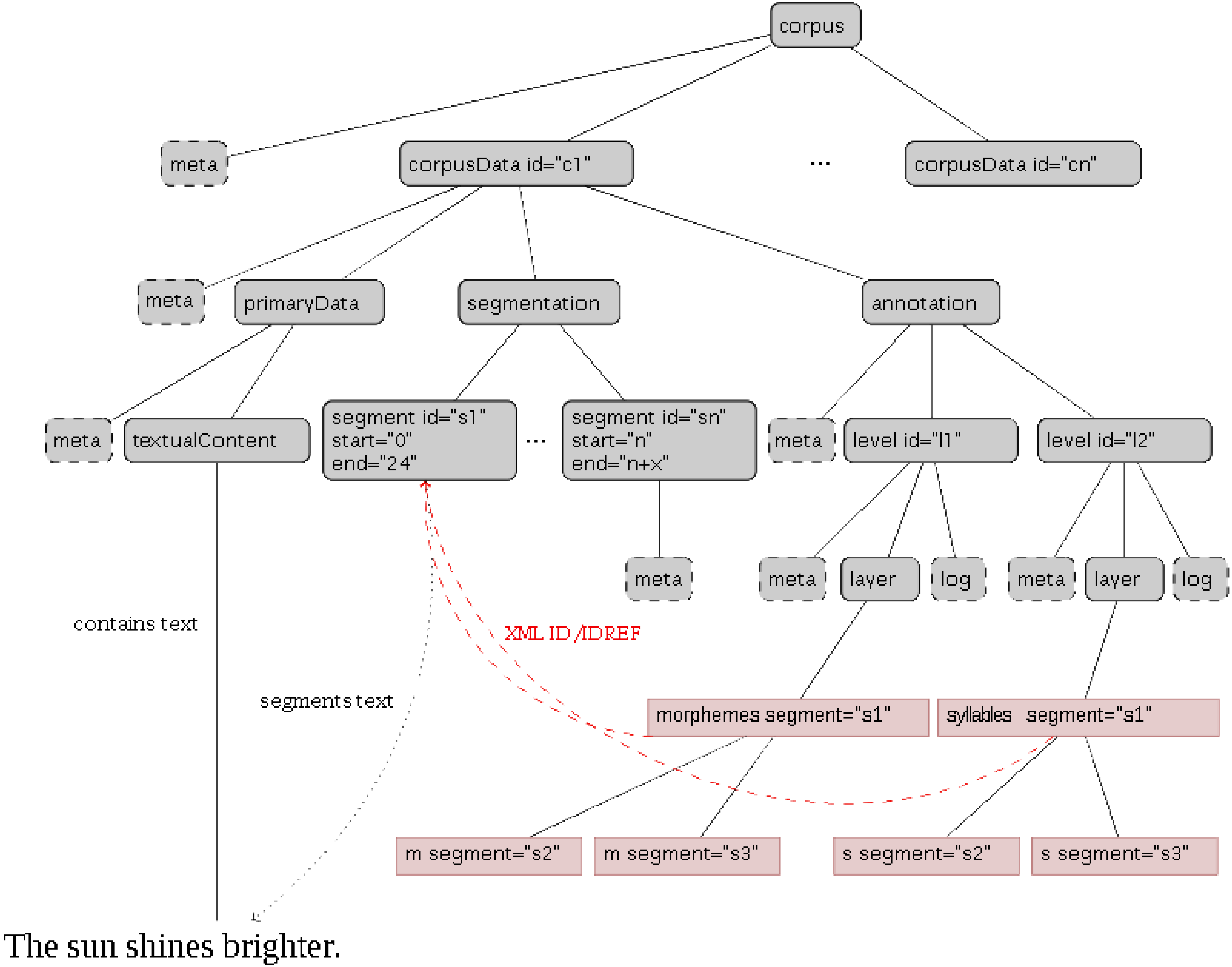
```
<?xml version="1.0" encoding="UTF-8"?>
<morphemes xmlns="http://www.xstandoff.net/morphemes">
  <morpheme>The</morpheme>
  <morpheme>sun</morpheme>
  <morpheme>shine</morpheme>
  <morpheme>s</morpheme>
  <morpheme>bright</morpheme>
  <morpheme>er</morpheme>.
</morphemes>
```

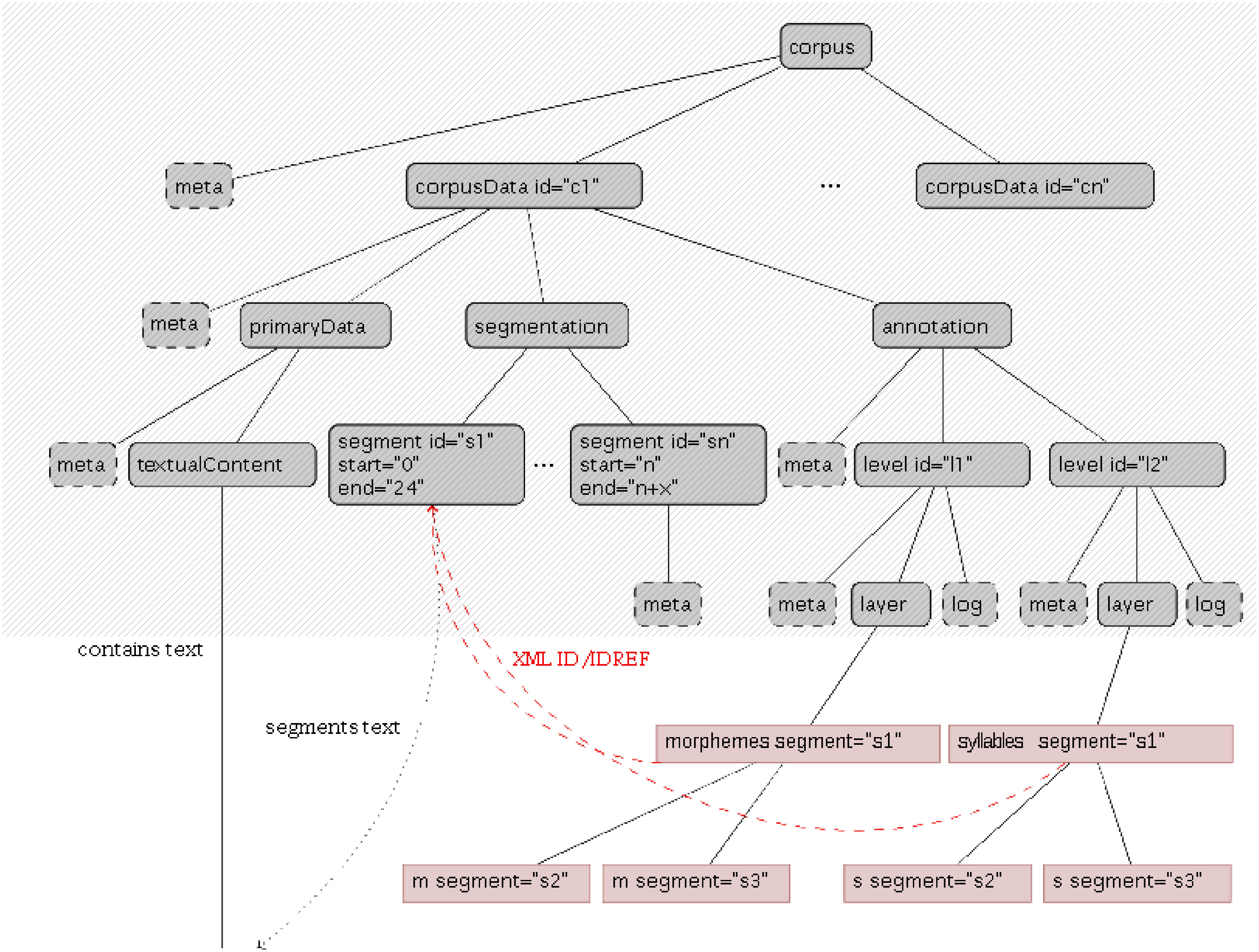
```
<?xml version="1.0" encoding="UTF-8"?>
<syllables xmlns="http://www.xstandoff.net/syllables">
  <syllable>The</syllable>
  <syllable>sun</syllable>
  <syllable>shines</syllable>
  <syllable>brigh</syllable>
  <syllable>ter</syllable>.
</syllables>
```

XStandoff – A simple example

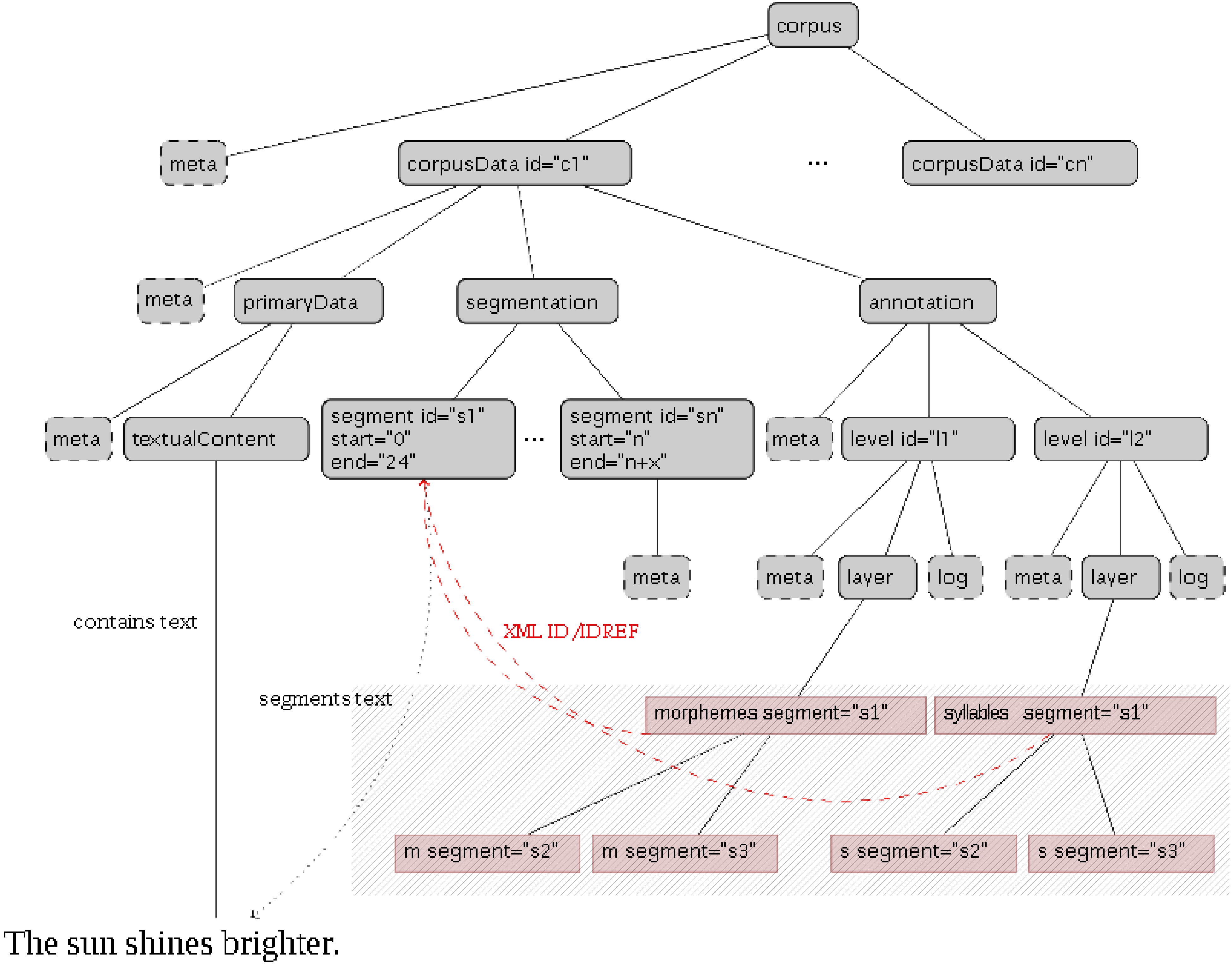
```
<?xml version="1.0" encoding="UTF-8"?>
<morphemes xmlns="http://www.xstandoff.net/morphemes">
  <morpheme>The</morpheme>
  <morpheme>sun</morpheme>
  <morpheme>shine</morpheme>
  <morpheme>s</morpheme>
  <morpheme>bright</morpheme>
  <morpheme>er</morpheme>.
</morphemes>

<?xml version="1.0" encoding="UTF-8"?>
<syllables xmlns="http://www.xstandoff.net/syllables">
  <syllable>The</syllable>
  <syllable>sun</syllable>
  <syllable>shines</syllable>
  <syllable>brigh</syllable>
  <syllable>ter</syllable>.
</syllables>
```


corpus
meta
corpusData id="c1"
...
corpusData id="cn"
meta
primaryData
segmentation
annotation
meta
textualContent
segment id="s1" start="0" end="24"
...
segment id="sn" start="n" end="n+x"
meta
level id="l1"
level id="l2"
meta
meta
layer
log
meta
layer
log
contains text
XML ID /IDREF
segments text
morphemes segment="s1"
syllables segment="s1"
m segment="s2"
m segment="s3"
s segment="s2"
s segment="s3"
The sun shines brighter.



The sun shines brighter.


corpus
meta
corpusData id="c1"
...
corpusData id="cn"
meta
primaryData
segmentation
annotation
meta
textualContent
segment id="s1"
start="0"
end="24"
...
segment id="sn"
start="n"
end="n+x"
meta
level id="l1"
level id="l2"
meta
meta
layer
log
meta
layer
log
contains text
XML ID /IDREF
segments text
morphemes segment="s1"
syllables segment="s1"
m segment="s2"
m segment="s3"
s segment="s2"
s segment="s3"
The sun shines brighter.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<xsf:corpusData xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.xstandoff.net/2009/xstandoff/1.1 http://www.xstandoff.net/2009/xstandoff/1.1/xsf.xsd"
  xmlns:xsf="http://www.xstandoff.net/2009/xstandoff/1.1" xml:id="c1" xsfVersion="1.1">
   <xsf:primaryData start="0" end="24" xml:lang="en">
      <xsf:textualContent>The sun shines brighter.</xsf:textualContent>
   </xsf:primaryData>
   <xsf:segmentation>
      <xsf:segment xml:id="seg1" type="char" start="0" end="24"/>
      <!-- [...] -->
      <xsf:segment xml:id="seg7" type="char" start="15" end="21"/>
      <xsf:segment xml:id="seg8" type="char" start="15" end="20"/>
      <xsf:segment xml:id="seg9" type="char" start="20" end="23"/>
      <xsf:segment xml:id="seg10" type="char" start="21" end="23"/>
   </xsf:segmentation>
   <xsf:annotation>
      <xsf:level xml:id="l_syll">
        <xsf:layer xmlns:morph="http://www.xstandoff.net/morphemes"
          xsi:schemaLocation="http://www.xstandoff.net/morphemes morphemes.xsd" priority="0">
            <morph:morphemes xsf:segment="seg1">
               <morph:morpheme xsf:segment="seg2"/>
               <morph:morpheme xsf:segment="seg3"/>
               <!-- [...] -->
               <morph:morpheme xsf:segment="seg10"/>
            </morph:morphemes>
         </xsf:layer>
      </xsf:level>
      <xsf:level xml:id="l_morph">
        <xsf:layer xmlns:syll="http://www.xstandoff.net/syllables"
          xsi:schemaLocation="http://www.xstandoff.net/syllables syllables.xsd" priority="1">
            <syll:syllables xsf:segment="seg1">
               <syll:syllable xsf:segment="seg2"/>
               <syll:syllable xsf:segment="seg3"/>
               <!-- [...] -->
               <syll:syllable xsf:segment="seg9"/>
            </syll:syllables>
         </xsf:layer>
      </xsf:level>
   </xsf:annotation>
</xsf:corpusData>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<xsf:corpusData xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xsi:schemaLocation="http://www.xstandoff.net/2009/xstandoff/1.1 http://www.xstandoff.net/2009/xstandoff/1.1/xsf.xsd"
 xmlns:xsf="http://www.xstandoff.net/2009/xstandoff/1.1" xml:id="c1" xsfVersion="1.1">
  <xsf:primaryData start="0" end="24" xml:lang="en">
    <xsf:textualContent>The sun shines brighter.</xsf:textualContent>
  </xsf:primaryData>
  <xsf:segmentation>
    <xsf:segment xml:id="seg1" type="char" start="0" end="24"/>
    <!-- [...] -->
    <xsf:segment xml:id="seg7" type="char" start="15" end="21"/>
    <xsf:segment xml:id="seg8" type="char" start="15" end="20"/>
    <xsf:segment xml:id="seg9" type="char" start="20" end="23"/>
    <xsf:segment xml:id="seg10" type="char" start="21" end="23"/>
  </xsf:segmentation>
  <xsf:annotation>
    <xsf:level xml:id="l_syll">
     <xsf:layer xmlns:morph="http://www.xstandoff.net/morphemes"
       xsi:schemaLocation="http://www.xstandoff.net/morphemes morphemes.xsd" priority="0">
        <morph:morphemes xsf:segment="seg1">
          <morph:morpheme xsf:segment="seg2"/>
          <morph:morpheme xsf:segment="seg3"/>
          <!-- [...] -->
          <morph:morpheme xsf:segment="seg10"/>
        </morph:morphemes>
      </xsf:layer>
    </xsf:level>
    <xsf:level xml:id="l_morph">
     <xsf:layer xmlns:syll="http://www.xstandoff.net/syllables"
       xsi:schemaLocation="http://www.xstandoff.net/syllables syllables.xsd" priority="1">
        <syll:syllables xsf:segment="seg1">
          <syll:syllable xsf:segment="seg2"/>
          <syll:syllable xsf:segment="seg3"/>
          <!-- [...] -->
          <syll:syllable xsf:segment="seg9"/>
        </syll:syllables>
      </xsf:layer>
    </xsf:level>
  </xsf:annotation>
</xsf:corpusData>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<xsf:corpusData xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.xstandoff.net/2009/xstandoff/1.1 http://www.xstandoff.net/2009/xstandoff/1.1/xsf.xsd"
  xmlns:xsf="http://www.xstandoff.net/2009/xstandoff/1.1" xml:id="c1" xsfVersion="1.1">
   <xsf:primaryData start="0" end="24" xml:lang="en">
      <xsf:textualContent>The sun shines brighter.</xsf:textualContent>
   </xsf:primaryData>
   <xsf:segmentation>
      <xsf:segment xml:id="seg1" type="char" start="0" end="24"/>
      <!-- [...] -->
      <xsf:segment xml:id="seg7" type="char" start="15" end="21"/>
      <xsf:segment xml:id="seg8" type="char" start="15" end="20"/>
      <xsf:segment xml:id="seg9" type="char" start="20" end="23"/>
      <xsf:segment xml:id="seg10" type="char" start="21" end="23"/>
   </xsf:segmentation>
   <xsf:annotation>
      <xsf:level xml:id="l_syll">
        <xsf:layer xmlns:morph="http://www.xstandoff.net/morphemes"
          xsi:schemaLocation="http://www.xstandoff.net/morphemes morphemes.xsd" priority="0">
            <morph:morphemes xsf:segment="seg1">
               <morph:morpheme xsf:segment="seg2"/>
               <morph:morpheme xsf:segment="seg3"/>
               <!-- [...] -->
               <morph:morpheme xsf:segment="seg10"/>
            </morph:morphemes>
         </xsf:layer>
      </xsf:level>
      <xsf:level xml:id="l_morph">
        <xsf:layer xmlns:syll="http://www.xstandoff.net/syllables"
          xsi:schemaLocation="http://www.xstandoff.net/syllables syllables.xsd" priority="1">
            <syll:syllables xsf:segment="seg1">
               <syll:syllable xsf:segment="seg2"/>
               <syll:syllable xsf:segment="seg3"/>
               <!-- [...] -->
               <syll:syllable xsf:segment="seg9"/>
            </syll:syllables>
         </xsf:layer>
      </xsf:level>
   </xsf:annotation>
</xsf:corpusData>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<xsf:corpusData xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.xstandoff.net/2009/xstandoff/1.1 http://www.xstandoff.net/2009/xstandoff/1.1/xsf.xsd"
  xmlns:xsf="http://www.xstandoff.net/2009/xstandoff/1.1" xml:id="c1" xsfVersion="1.1">
   <xsf:primaryData start="0" end="24" xml:lang="en">
      <xsf:textualContent>The sun shines brighter.</xsf:textualContent>
   </xsf:primaryData>
   <xsf:segmentation>
      <xsf:segment xml:id="seg1" type="char" start="0" end="24"/>
      <!-- [...] -->
      <xsf:segment xml:id="seg7" type="char" start="15" end="21"/>
      <xsf:segment xml:id="seg8" type="char" start="15" end="20"/>
      <xsf:segment xml:id="seg9" type="char" start="20" end="23"/>
      <xsf:segment xml:id="seg10" type="char" start="21" end="23"/>
   </xsf:segmentation>
   <xsf:annotation>
      <xsf:level xml:id="l_syll">
        <xsf:layer xmlns:morph="http://www.xstandoff.net/morphemes"
          xsi:schemaLocation="http://www.xstandoff.net/morphemes morphemes.xsd" priority="0">
            <morph:morphemes xsf:segment="seg1">
               <morph:morpheme xsf:segment="seg2"/>
               <morph:morpheme xsf:segment="seg3"/>
               <!-- [...] -->
               <morph:morpheme xsf:segment="seg10"/>
            </morph:morphemes>
         </xsf:layer>
      </xsf:level>
      <xsf:level xml:id="l_morph">
        <xsf:layer xmlns:syll="http://www.xstandoff.net/syllables"
          xsi:schemaLocation="http://www.xstandoff.net/syllables syllables.xsd" priority="1">
            <syll:syllables xsf:segment="seg1">
               <syll:syllable xsf:segment="seg2"/>
               <syll:syllable xsf:segment="seg3"/>
               <!-- [...] -->
               <syll:syllable xsf:segment="seg9"/>
            </syll:syllables>
         </xsf:layer>
      </xsf:level>
   </xsf:annotation>
</xsf:corpusData>
```

XStandoff and (x)LMNL

- XStandoff extends the limits of XML's data model but otherwise remains in the XML world
- Similar to that is LMNL's (Layered Markup and Annotation Language) XML serialisation by Tennision introduced
- LMNL addresses the following XML components:
  - Syntax (when using the sawtooth syntax)
  - Data model (Ranges instead of hierarchies)
  - Validation (depending on the syntax used)

LMNL

LMNL differs from XML

- LMNL uses (possibly overlapping) ranges (sequences of characters) instead of hierarchies
- Ranges can be annotated
- LMNL supports structured annotations (annotations can be miniature documents – think of structured attributes in plain XML)
- In contrast to XML attributes LMNL annotations are ordered and more than one annotation of the same name and range can be given

LMNL – A simple example

```
[excerpt [source [book}1 Kings{book] [chapter}12{chapter]]}
[verse}And he said unto them, [q}What counsel give ye that we may answer this
people, who have spoken to me, saying, [q}Make the yoke which thy father did
put upon us lighter?{q]{q]{verse]
[verse}And the young men that were grown up with him spake unto him, saying,
[q}Thus shalt thou speak unto this people that spake unto thee, saying, [q=i}Thy
father made our yoke heavy, but make thou it lighter unto us;{q=i] thus shalt
thou say unto them, [q=j}My little finger shall be thicker than my father's loins.
{verse]
[verse}And now whereas my father did lade you with a heavy yoke, I will add to
your yoke: my father hath chastised you with whips, but I will chastise you with
scorpions.{q=j]{q]{verse]
[verse}So Jeroboam and all the people came to Rehoboam the third day, as the
king had appointed, saying, [q}Come to me again the third day.{q]{verse]
{excerpt]
```

```
<?xml version="1.0" encoding="UTF-8"?>
<x:document xmlns:x="http://lmnl-markup.org/ns/xLMNL" ID="N.d2821e1"
  base-uri="file:/homei/users/wapiez/courseweb_html/cocoon_pub/LMNL/lmnl/1kings12.lmnl">
  <x:content>
    <x:span start="0" end="1" layer="N.d2821e1" ranges="R.d2821e2"> </x:span>
    <x:span start="1" end="24" layer="N.d2821e1" ranges="R.d2821e2 R.d2821e12">And he said unto
      them, </x:span>
    <x:span start="24" end="108" layer="N.d2821e1" ranges="R.d2821e2 R.d2821e12 R.d2821e15">What
      counsel give ye that we may answer this people, who have spoken to me, saying, </x:span>
    <!-- [...] -->
    <x:span start="752" end="753" layer="N.d2821e1"> </x:span>
  </x:content>
  <x:range start="0" end="752" ID="R.d2821e2" sl="1" so="1" name="excerpt" el="6" eo="9">
    <x:annotation ID="N.d2821e3" sl="1" so="10" el="1" eo="58" name="source">
      <x:annotation ID="N.d2821e4" sl="1" so="18" el="1" eo="36" name="book">
        <x:content>
          <x:span start="0" end="7" layer="N.d2821e4">1 Kings</x:span>
        </x:content>
      </x:annotation>
      <x:annotation ID="N.d2821e7" sl="1" so="38" el="1" eo="57" name="chapter">
        <x:content>
          <x:span start="0" end="2" layer="N.d2821e7">12</x:span>
        </x:content>
      </x:annotation>
      <x:content/>
    </x:annotation>
  </x:range>
  <x:range start="1" end="163" ID="R.d2821e12" sl="2" so="1" name="verse" el="2" eo="188"/>
  <x:range start="24" end="163" ID="R.d2821e15" sl="2" so="31" name="q" el="2" eo="181"/>
  <!-- [...] -->
  <x:range start="720" end="751" ID="R.d2821e57" sl="5" so="106" name="q" el="5" eo="142"/>
</x:document>
```

```xml
<?xml version="1.0" encoding="UTF-8"?>
<x:document xmlns:x="http://lmnl-markup.org/ns/xLMNL" ID="N.d2821e1"
  base-uri="file:/homei/users/wapiez/courseweb_html/cocoon_pub/LMNL/lmnl/1kings12.lmnl">
  <x:content>
    <x:span start="0" end="1" layer="N.d2821e1" ranges="R.d2821e2"> </x:span>
    <x:span start="1" end="24" layer="N.d2821e1" ranges="R.d2821e2 R.d2821e12">And he said unto
      them, </x:span>
    <x:span start="24" end="108" layer="N.d2821e1" ranges="R.d2821e2 R.d2821e12 R.d2821e15">What
      counsel give ye that we may answer this people, who have spoken to me, saying, </x:span>
    <!-- [...] -->
    <x:span start="752" end="753" layer="N.d2821e1"> </x:span>
  </x:content>
  <x:range start="0" end="752" ID="R.d2821e2" sl="1" so="1" name="excerpt" el="6" eo="9">
    <x:annotation ID="N.d2821e3" sl="1" so="10" el="1" eo="58" name="source">
      <x:annotation ID="N.d2821e4" sl="1" so="18" el="1" eo="36" name="book">
        <x:content>
          <x:span start="0" end="7" layer="N.d2821e4">1 Kings</x:span>
        </x:content>
      </x:annotation>
      <x:annotation ID="N.d2821e7" sl="1" so="38" el="1" eo="57" name="chapter">
        <x:content>
          <x:span start="0" end="2" layer="N.d2821e7">12</x:span>
        </x:content>
      </x:annotation>
      <x:content/>
    </x:annotation>
  </x:range>
  <x:range start="1" end="163" ID="R.d2821e12" sl="2" so="1" name="verse" el="2" eo="188"/>
  <x:range start="24" end="163" ID="R.d2821e15" sl="2" so="31" name="q" el="2" eo="181"/>
  <!-- [...] -->
  <x:range start="720" end="751" ID="R.d2821e57" sl="5" so="106" name="q" el="5" eo="142"/>
</x:document>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<x:document xmlns:x="http://lmnl-markup.org/ns/xLMNL" ID="N.d2821e1"
  base-uri="file:/homei/users/wapiez/courseweb_html/cocoon_pub/LMNL/lmnl/1kings12.lmnl">
  <x:content>
    <x:span start="0" end="1" layer="N.d2821e1" ranges="R.d2821e2"> </x:span>
    <x:span start="1" end="24" layer="N.d2821e1" ranges="R.d2821e2 R.d2821e12">And he said unto
      them, </x:span>
    <x:span start="24" end="108" layer="N.d2821e1" ranges="R.d2821e2 R.d2821e12 R.d2821e15">What
      counsel give ye that we may answer this people, who have spoken to me, saying, </x:span>
    <!-- [...] -->
    <x:span start="752" end="753" layer="N.d2821e1"> </x:span>
  </x:content>
  <x:range start="0" end="752" ID="R.d2821e2" sl="1" so="1" name="excerpt" el="6" eo="9">
    <x:annotation ID="N.d2821e3" sl="1" so="10" el="1" eo="58" name="source">
      <x:annotation ID="N.d2821e4" sl="1" so="18" el="1" eo="36" name="book">
        <x:content>
          <x:span start="0" end="7" layer="N.d2821e4">1 Kings</x:span>
        </x:content>
      </x:annotation>
      <x:annotation ID="N.d2821e7" sl="1" so="38" el="1" eo="57" name="chapter">
        <x:content>
          <x:span start="0" end="2" layer="N.d2821e7">12</x:span>
        </x:content>
      </x:annotation>
      <x:content/>
    </x:annotation>
  </x:range>
  <x:range start="1" end="163" ID="R.d2821e12" sl="2" so="1" name="verse" el="2" eo="188"/>
  <x:range start="24" end="163" ID="R.d2821e15" sl="2" so="31" name="q" el="2" eo="181"/>
  <!-- [...] -->
  <x:range start="720" end="751" ID="R.d2821e57" sl="5" so="106" name="q" el="5" eo="142"/>
</x:document>
```

LMNL

- LMNL shows some interesting extensions to XML's data model – and a possible linearisation
- Since it lacks support for hierarchies it abandons one of XML's data model key features

FtanML

- FtanML was developed during a summer academy by Michael Kay and others
- It addresses the following of XML's components:
  - Data model
  - Notation format
  - Constraint language
- The data model is similar to JSON's array of key-value pairs but adds elements, attributes and contents
- In contrast to JSON it supports XML's mixed content

FtanML

Some examples:

- JSON-like arrays

  [1, 2, “abc”, [1, 2]]

- XML-like FtanML element

  <para|Here is some <b|bold> text>

FtanML

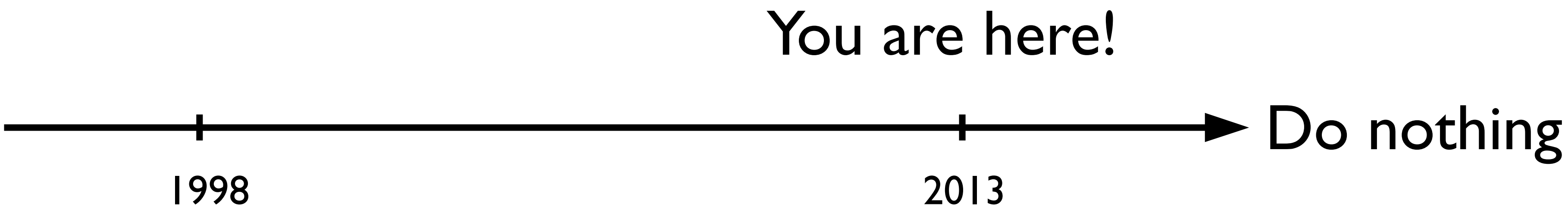
- There is a parser for FtanML and serialization for its object model in JSON and XML but no traversal language or validation language
- Up to now, FtanML is only a proposal – but there will be further development (hint: wait for this year's Balisage)

Quo vadis XML?

Quo vadis XML?

There are different development options:

- If we change nothing we may store any information we want – but it may be cumbersome (think of XStandoff and xLMNL) – but we'll benefit from XML's ecosystem



You are here!

Do nothing

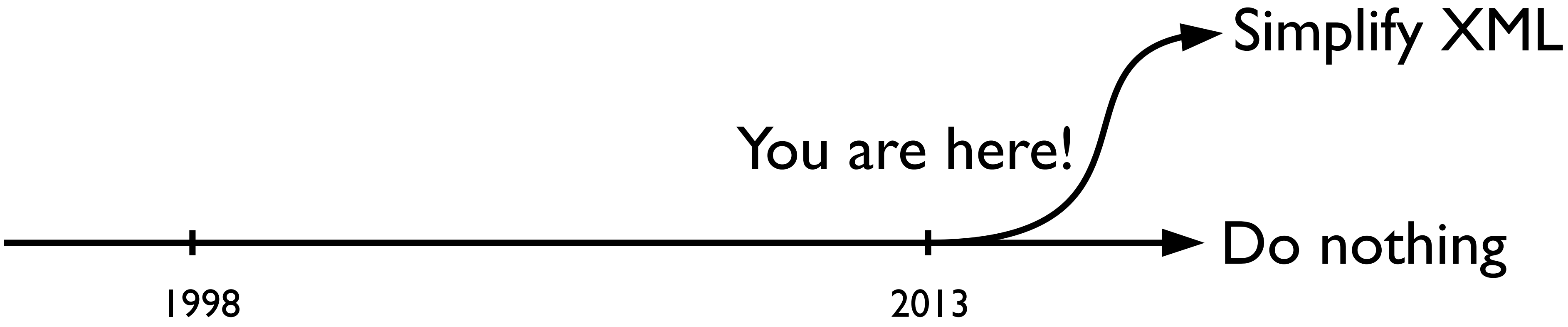
1998

2013

Quo vadis XML?

There are different development options:

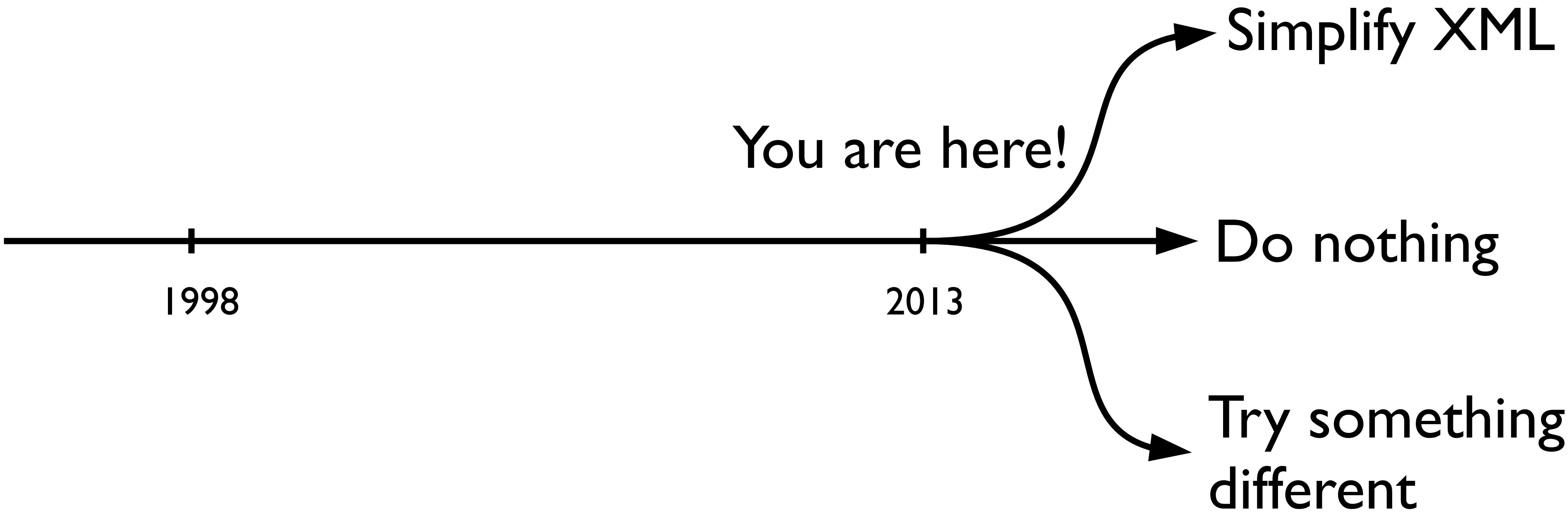
- We could go the MicroXML way and both simplify XML and adapt its data model



Quo vadis XML?

There are different development options:

- We can invent another meta language which a certain feature set like LMNL or FtanML – but we may lose the benefits of XML's ecosystem



I myself cannot answer which way is the best, but I could raise concerns for some points...

What I'd like to keep from current XML

- Hierarchical data model
- Support for mixed content models
- Processable with XML's toolchain (this does not have to mean XML's syntax, since XML's toolchain is capable of processing non-XML input, see Piez' LMNL sawtooth parser written in XSLT 2.0)
- Powerful constraint languages such as XSD 1.1 or RNG
- Powerful transformation and query languages such as XSLT and XQuery

What I'd like to see added

- Support for overlapping hierarchies and discontinuous elements
- Use a modular approach – simple for those that need only a fraction of its power and complex for those that need everything

<Happy_birthday_XML_nonetheless/>

Thank you for your attention!
Questions? Comments?

maik.stuehrenberg@uni-bielefeld.de
@mstuehrenberg

References

- All cliparts except the 'Pulp fiction' by http://openclipart.org (Public Domain license)
- 'Pulp fiction' SVG conversion taken from http://1.bp.blogspot.com/-mljch6QirOY/TbbQx4JIQ6I/AAAAAAAABPo/a_ui9B1BnvE/s320/Pulp_Fiction_Wallpaper_by_Oshowom.jpg
- Witt, Andreas (2004). Multiple hierarchies: New Aspects of an Old Solution. In: Proceedings of Extreme Markup Languages, Montréal
- Kay, Michael (2012). FtanML - A new markup language. http://dev.saxonica.com/blog/mike/2012/08/ #000193
- Piez, Wendell (2004). Half-steps toward LMNL. In: Proceedings of Extreme Markup Languages 2004, Montréal
- Piez, Wendell (2012). Luminescent: Parsing LMNL by XSLT Upconversion. In Proceedings of Balisage: The Markup Conference, Balisage Series on Markup Technologies vol. 8, Montréal
- Sperberg-McQueen, C. M., and Claus Huitfeldt (1999). Concurrent Document Hierarchies in MECS and SGML. In: Literary and Linguistic Computing 14.1:29–42
- Huitfeldt, Claus, and C. M. Sperberg-McQueen (2004). Markup Languages and Complex Documents (MLDC). Presentation given at the colloquium of the HKI (Historisch Kulturwissenschaftliche Informationsverarbeitung) of the University of Cologne, 10 Decembre 2004. http://old.hki.uni-koeln.de/material/huitfeldt.ppt
- Stührenberg, Maik and Daniela Goecke (2008). SGF – an integrated model for multiple annotations and its application in a linguistic domain. In: Proceedings of Balisage: The Markup Conference, Balisage Series on Markup Technologies, vol 1, Montréal
- Stührenberg, Maik and Daniel Jettka (2009). A toolkit for multi-dimensional markup: The development of SGF to XStandoff. In: Proceedings of Balisage: The Markup Conference 2009. Balisage Series on Markup Technologies, vol. 3, Montréal
- Tennison, Jeni (2002). Layered Markup and Annotation Language (LMNL). In: Proceedings of Extreme Markup Languages 2002, Montréal