



TEI P5における スキーマの生成・TEIとXSLのカスタマイズ

Espen S. Ore

Unit for Digital Documentation

Dept. of Linguistics and Scandinavian Studies

University of Oslo, Norway

(翻訳・加筆修正: 一般財団法人人文情報学研究所・永崎研宣)

はじめのはじめに

- この資料は、oXygenを用いてTEI/XMLに準拠した文書を作成する方法を理解するためのチュートリアルです。
- oXygenは事前にインストールしておいてください。
- まず、oXygenでTEI P5のスキーマを用いたファイル作成を行い、
- oXygenによる検証 (verification) の便利さを確認し、
- Romaを使ったスキーマのカスタマイズの方法を学びます。
- その後、XSLの使い方の基本を学びます。
- XML、スキーマ、XSLについては、XMLの初歩のガイドブックならどれにでも解説してありますのでそこまでは自分で学習してください。

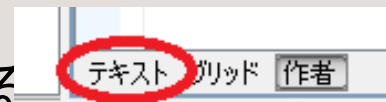
はじめに

- まずはシンプルなTEIスキーマからはじめましょう。
- ブラウザで以下のURLにアクセスしてください。
- <http://www.tei-c.org/Roma/>
- デフォルトのままで「start」をクリックしてください。
- デフォルトの値はそのままにして、トップメニューの「Schema」タブをクリックしてください。
- デフォルトの値はそのまま、「Generate」をクリックしてください。
- .rncという拡張子のファイルをハードディスクに保存してください。（どこに保存したかも忘れずに）
- Romaで「Save customization」を選んで、.xmlという拡張子のファイルを保存してください。
- （Romaのページはそのまま開いておいてください。）

TEI文書の作成

- Oxygenで、「ファイル」メニューから「新規作成」を選んでください。
- ダイアログ中の「新しい文書」から「XML文書」を選んで、ダイアログ下部の「カスタマイズ」をクリックしてください。
- 次に「Customize editor」というダイアログが表示されますので、スキーマURL: のところで、先ほどRomaからダウンロードしたスキーマを指定してください。
- 「スキーマ種別」のところに、自動的に「Relax NG コンパクトシンタックス」となるはずですので確認してから、「作成」をクリックしてください。
- 新規ファイルが開きますので、次に、画面下部に並んでいる「テキスト グリッド 作者」という3つのボタンの中の「テキスト」をクリックしてください。

- 開いた文書のタグに赤い下線がついているタグがあります。これらは、スキーマ上、誤った記述になっていることを示しており、画面下部にエラーの理由が表示されています。これらは、記述を修正してスキーマに準拠する形になれば消えます。以下、スキーマに準拠するように文書を修正していきます。



正しいteiHeader を作ってみましょう。

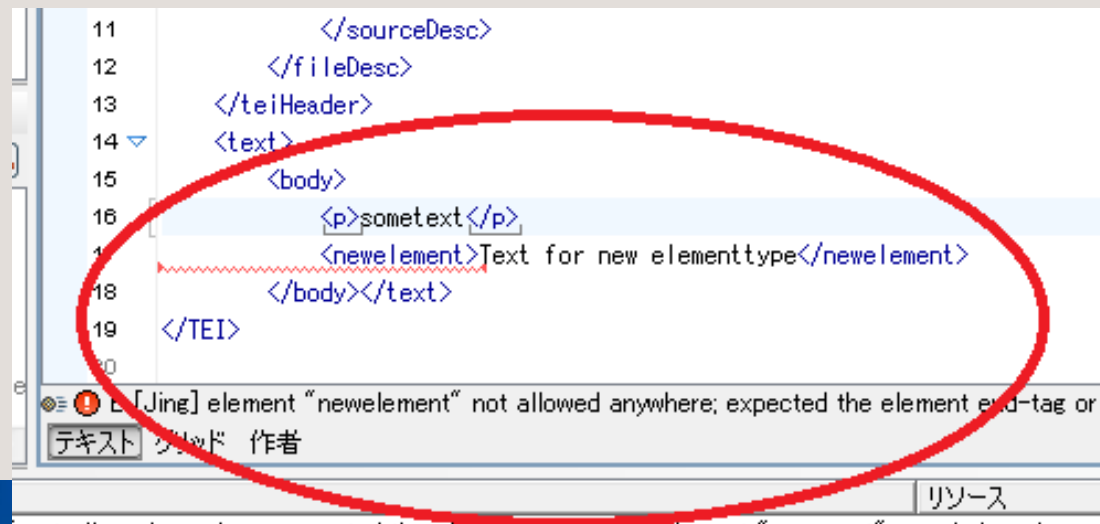
文書(の形式)が(スキーマに対して)妥当となるように必要なヘッダエレメントを付け加えて、TEIヘッダを編集してみてください。なお、このヘッダは、すべてのTEI文書に必要となるものです。

(TEIでは、「文書に関する様々な説明」をヘッダに付けておくためのルールがあります。このルールに従うことで、文書に関する様々な付帯情報をやりとりしようとしています。)

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title></title>
    </titleStmt>
    <publicationStmt><p>A short publicationstatement</p></publicationStmt>
    <sourceDesc><p>and another short header element</p></sourceDesc>
  </fileDesc>
</teiHeader>
```

文書にテキスト本文を追加する

- `<text>`エレメントの中に`<body>`エレメントを一つ追加してください。
- `<body>`エレメントの中で、まずは`<p>sometext</p>`というエレメントを作ってみて、それから、以下のエレメントを付け加えてください。ここまでで、赤い下線はすべて消えるはずですが。
- 次に、以下の文章を画像の例のように追記してみましょう。`<newelement> Text for new elementtype</newelement>`
- この場合はエラーが表示されます。このエレメントはオリジナルのものだからです。これをエラーなしで使えるようにするにはスキーマをカスタマイズする必要があります。



```
11         </sourceDesc>
12     </fileDesc>
13 </teiHeader>
14 <text>
15     <body>
16         <p>sometext</p>
17         <newelement>Text for new elementtype</newelement>
18     </body></text>
19 </TEI>
```

エラーメッセージ: [Jing] element "newelement" not allowed anywhere; expected the element end-tag or

スキーマの変更

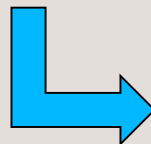
- スキーマのカスタマイズは、技術的にかなりややこしく、人文学者が取り組むにはかなり難しいです。そこで、TEIでは、それを容易にできるWebアプリケーションを用意しており、ここではそれを試してみます。
- 先ほどのRomaの画面で一番左の「New」をクリックしてください。(最初の画面に戻ります。)
- 「"Open existing customization"」を選び、最初に保存したxmlファイルを選択してから「start」をクリックしてください。……(A)
- 「Add Elements」タブを選んでから、
 - 「Name」のところに「newelement」と入力し
 - 「model」の一覧から「model.common」をチェックして、
 - 一番下のセレクトメニューから「Text」を選択してから、
 - 「save」をクリックしてください。
- Webページの上部に「 Create new element New element newelement successfully created 」と表示されるはずですが。
- 最初の手順と同様にして、「schema」タブをクリックして画面下部の「Generate」をクリックしてファイルを保存ください。(さっきとは違うファイル名で保存！)
- Romaの「Customize」タブを選んで「Namespace for new elements」の行をコピーしておいてください。(例 : <http://www.example.org/ns/nonTEI>)
- では、Oxygenに戻りましょう。

スキーマの変更 2



- 現在のスキーマを参照している行を削除してください。(例: `<?xml-model href="myTEI.rnc" type="application/relax-ng-compact-syntax"?>`)
 - ファイルメニューの「文書」からサブメニューの「スキーマ」を選んでから「スキーマの割り当て」を選び、そこで、「RelaxNG Schema」タブが自動選択されていることを確認し、先ほど保存した新しいスキーマファイルを選択してください。
 - ルートエレメント (`<TEI ...>`) の「xmlns="http://www.tei-c.org/ns/1.0"」という属性・値の組の後に、一つ半角スペースを入れてから
「xmlns:nonTEI="先ほどコピーしたNamespace for new elementsの行"」
を追加してください。
- xmlns については <http://www.kanzaki.com/docs/sw/names.html> などを参照。
- 文書中の `<newelement>` を `<nonTEI:neuelement>` と書き換えてください。(終了タグは自動的に変更されるはずなのでそのままにしておいてください。)

これでエラーが消えるはずです。



```
14 <text>
15   <body>
16     <p>somertext</p>
17     <nonTEI:neuelement>Text for new elementtype</nonTEI:neuelement>
18   </body></text>
19 </TEI>
20
```

テキスト グリッド 作者

HWS2010#emimin.xml

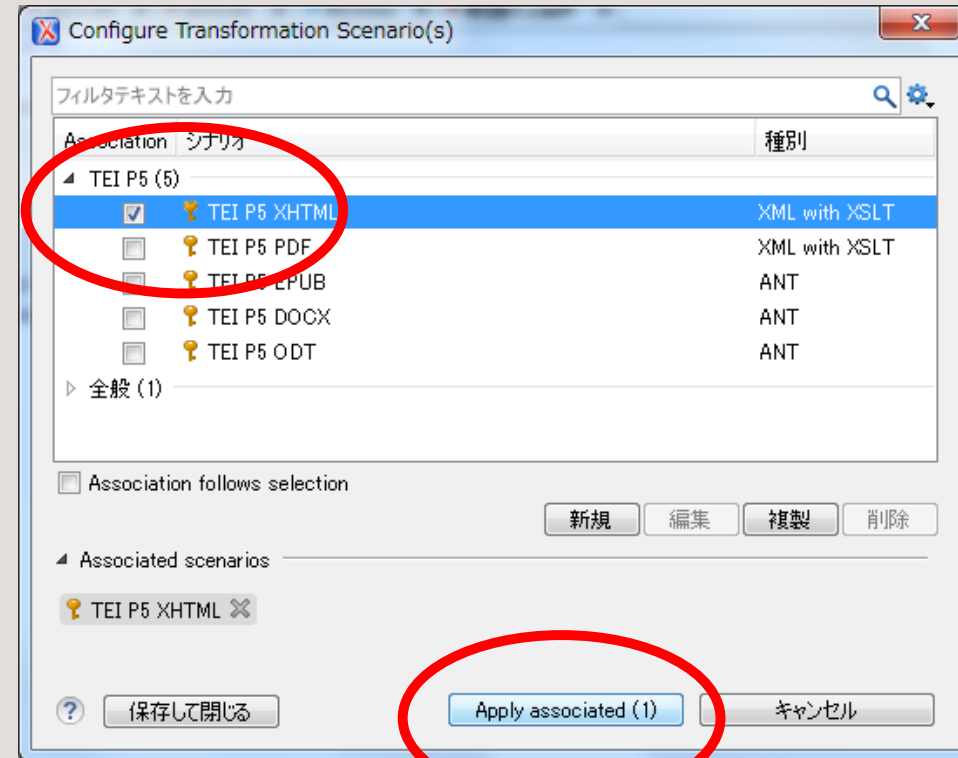
ドキュメントは妥当です。

閑話休題

- このように、TEIでは、TEIガイドラインにとって未知な資料にも適切に対応できるように、ユーザに対して、スキーマをカスタマイズするための手立てを公式に提供しています。
- また、カスタマイズした内容が、ガイドラインとして採り上げられるべきものであると考えた場合、TEIのユーザメーリングリストに提起することでその道筋が開けるようになっています。
- TEIガイドラインは、多様な資料に対応していくべく、必要に応じたバージョンアップを重ねています。
- ここから先は、TEIの目次を見ながらご自身にあったタグ・エレメントの使い方を検討してみてください。分野や資料によって、かなり異なったマークアップの手法が提示されています。

テキストの整形: XSLのデフォルト

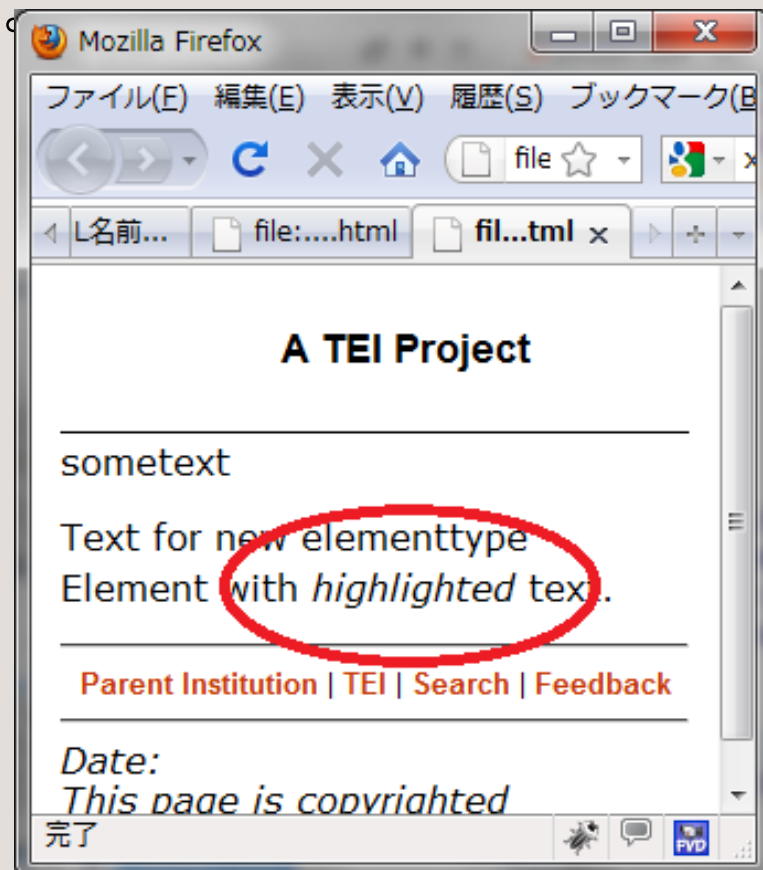
- Oxygenで、「文書」メニューから「変換」→「変換シナリオの設定」を選んでください。
- 以下のウインドウから、「TEI P5 XHTML」にチェックを入れてから「変換する(Apply associated)」ボタンをクリックしてください。



デフォルト設定のWebブラウザが開いてこんな感じに表示されるはずですが。

テキストの整形: XSLのデフォルト

- 次に、Oxygenで、<newelement>の行の次に
<p>Element with <hi rend="italic">highlighted</hi> text.</p>
と入力してから先ほどの手順で変換してみてください。
- すると、このように、イタリックで表示されます。
- 次に、italic の後ろに半角スペースを一つ入れて
から bold と追記し、先ほどの手順でもう一度変換
してみてください。そうすると、イタリックかつボールド
になるはずです。





XSLを一から書いてみましょう(1)

- Oxygenで「ファイル」→「新規作成」を選んでから、「xsl stylesheet」を選んで、次に「バージョン2.0」を選んでください。
- 「version =" 2.0" 」の後に以下のものを追加してください。

```
xpath-default-namespace="http://www.tei-c.org/ns/1.0"  
xmlns:nonTEI="http://www.example.org/ns/nonTEI"
```

XSLを一から書いてみましょう(2)

- 終了タグ</stylesheet>(もしくは</xsl:stylesheet>)の前に以下のものを追加してください。

```
<xsl:output method="html"/>
  <xsl:template match="/">
    <html>
      <head><title>Test document</title></head>
      <body>
        <xsl:apply-templates/>
      </body>
    </html>
  </xsl:template>
```

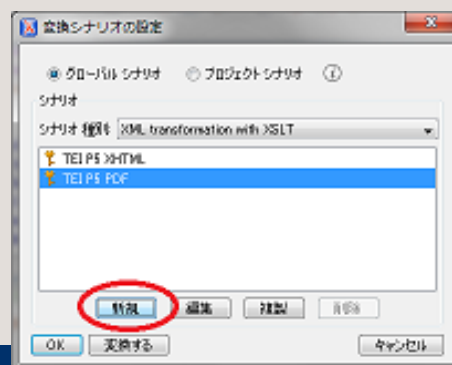
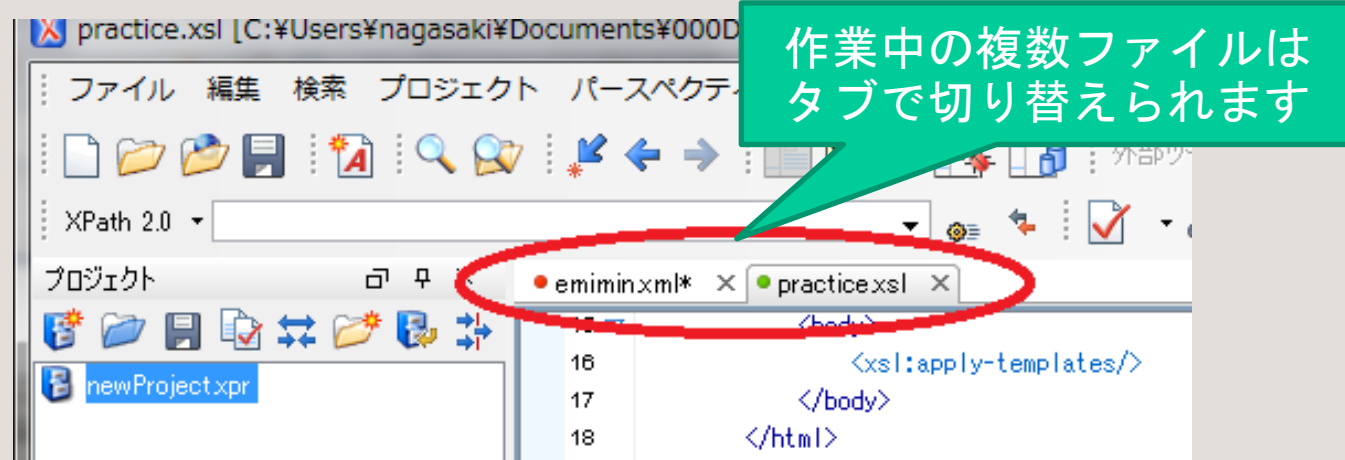
XSLを一から書いてみましょう(3)

- 続けて、終了タグ</stylesheet>(もしくは</xsl:stylesheet>)の前に以下のものも追加してください。

```
<xsl:template match="div|p">
  <xsl:apply-templates/>
</xsl:template>
<xsl:template match="hi">
  <i>
    <xsl:apply-templates/>
  </i>
</xsl:template>
<xsl:template match="nonTEI:newelement">
  <b>
    <xsl:apply-templates/>
  </b>
</xsl:template>
```

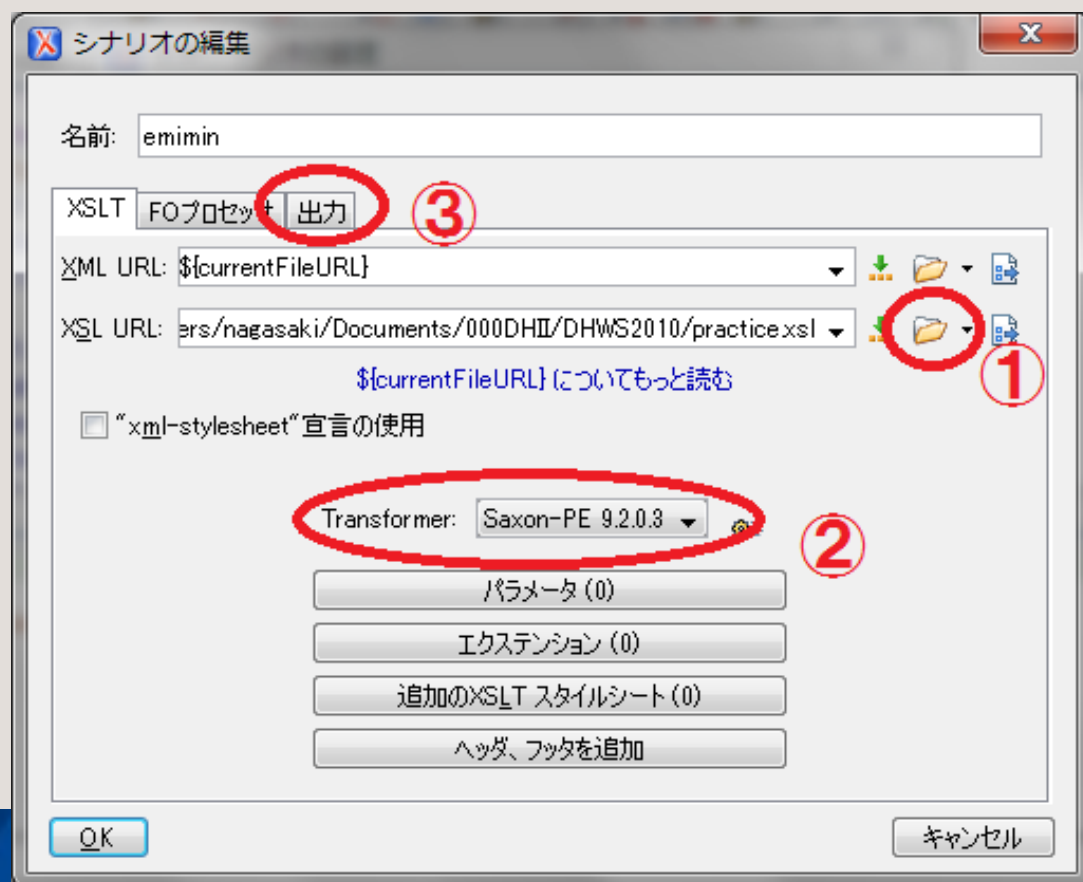
XSLを一から書いてみましょう(4)-1

- このxslファイルを保存してください。(保存した場所とファイル名を忘れないようにしてください。)
- Oxygenで、さきほど作成したxmlファイルのタブを選択して、xmlファイルを開いてください。
- 次に、さきほどと同様に、「文書」メニューから「変換」→「変換シナリオの設定」を選んでください。
- 「変換シナリオの設定」ダイアログで「新規」ボタンをクリックしてください。



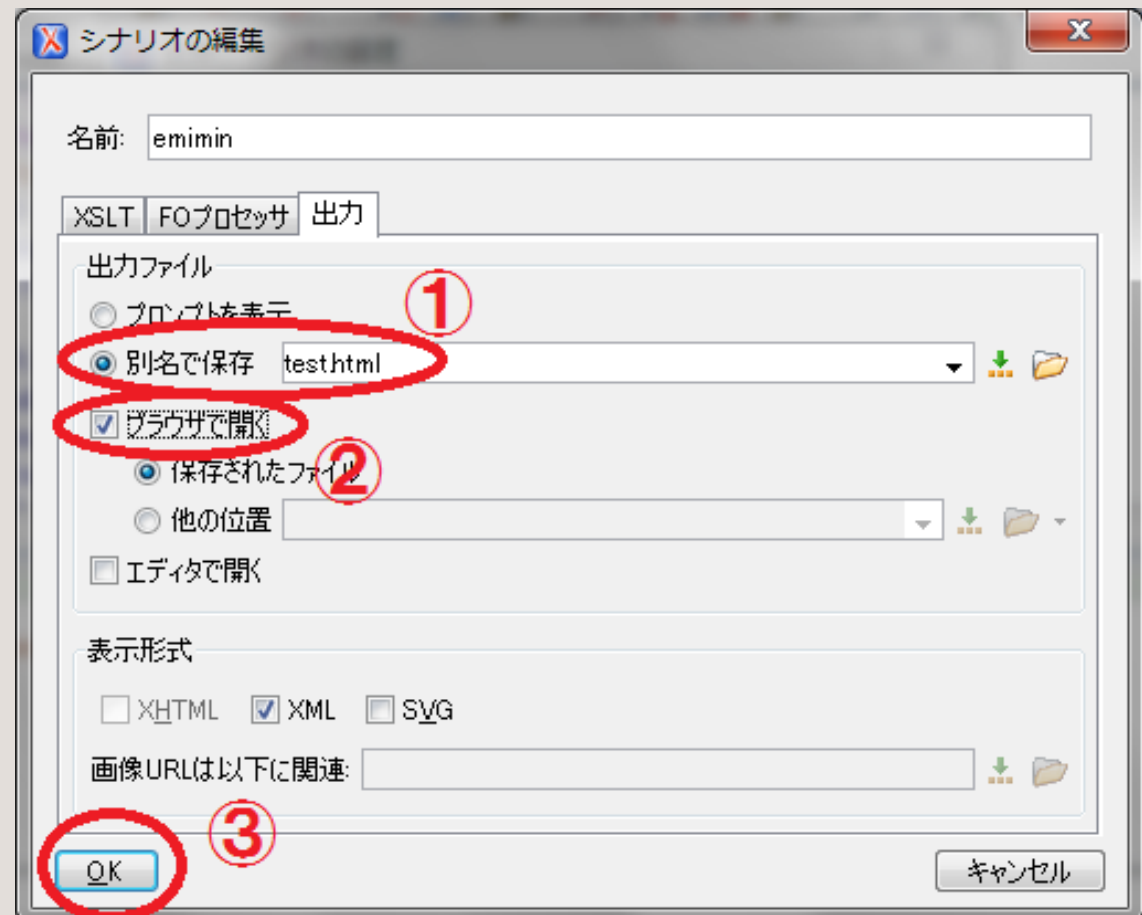
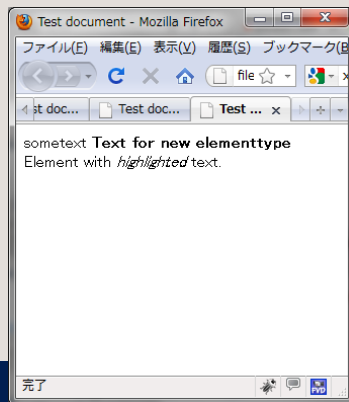
XSLを一から書いてみましょう(4)-2

- そうすると、「シナリオの編集」ウィンドウが開きますので、先ほど保存したXSLファイルをXSL URLのところ指定してください。(この行の右側のフォルダアイコンをクリックするとファイル名の指定ができます)。
- Transformer:のところ「 Saxon PE 9 .x」を選んでください。
- 次に「出力」タブをクリックして
- ください。



XSLを一から書いてみましょう(5)

- 「別名で保存」のところに「test.html」と入力してください。
- 「ブラウザで開く」にチェックを入れてください。
- 次に、「OK」をクリックすると、「変換シナリオの設定」ダイアログに戻ります。ここで、「変換する」ボタンをクリックしてください。
- そうすると先ほど作成したXSLファイルによる変換が行われ、Webブラウザに表示されます。



XSLを一から書いてみましょう(6)

- Oxygen上で、先ほどのxslファイルに戻ってください。(タブの切り替えだけでできるはずです。)
- `</xsl:stylesheet>`の前に以下のものを追加してください。

```
<xsl:template match="title">
  <h1>
    <xsl:apply-templates/>
  </h1>
</xsl:template>
```

- このxslファイルを保存してください。
- 先ほどのxmlファイルに戻ってください。(タブの切り替えで)
- 先ほどの手順で「変換」を行ってみてください。