

Tag2Xml-Konverter

Unter Tag2Xml-Konverter verstehen wir den Algorithmus der die neuen Tags in valides XML umrechnet.
Also z.B. aus

```
(<POS: N.>) {<Dom.: Bot.>} Blumenkohl<Gen.: m>; Karfiol<Gen.: m> (<Etym.: von engl. <For.: cauliflower>;  
<Scientif.: Brassica oleracea var. botrytis>).
```

wird:

```
<entry id="4626535">  
  <etym lang="en" origin="cauliflower">von engl. "cauliflower"</etym>  
  <gramGrp>  
    <pos type="N"/>  
  </gramGrp>  
  <usg type="dom">Bot</usg>  
  <sense>  
    <trans>  
      <tr>Blumenkohl</tr>  
    </trans>  
    <trans lang="scientific">  
      <tr>Brassica oleracea var. botrytis</tr>  
    </trans>  
    <trans>  
      <tr>Karfiol</tr>  
    </trans>  
  </sense>  
</entry>
```

Dieser Algorithmus kann auf dem Testrechner getestet werden. Feedback bitte hier oder ins Forum.
Je besser dieser Algorithmus ist, desto besser wird unser neuer Index und damit die neue Suche.

Für Entwickler

Wie kann der Tag2XmlConverter einfach getestet werden?

1. Den Home-Pfad in der dem /webroot/WEB-INF/classes/wadoku.properties" anpassen. Er muss auf den Projekt-Ordner verweisen. Ohne Slash am Ende. also z.B:

```
home=c:/data/wadoku/dev/website
```

2. Um zu bestimmen welchen getaggten Wadoku-Eintrag er ins XML konvertieren soll, kann man in der /webroot/xml/entry.txt einfach vor einer beliebigen Zeile ein "x" machen. Also z.B.:

```
x2||(<POS: N.>) blauer Himmel<Gen.: m>; Himmelsbläue<Gen.: f>.
```

3. Dann die Main-Methode aufrufen

Die Ausgabe sollte dann etwa sein:

```
-----  
(<POS: N.>) blauer Himmel<Gen.: m>; Himmelsbläue<Gen.: f>.
```

```
-----  
<?xml version="1.0" encoding="UTF-8"?>  
<entry id="2">  
  <gramGrp>  
    <pos type="N"/>  
  </gramGrp>  
  <sense>  
    <trans>  
      <tr>blauer Himmel</tr>  
    </trans>  
    <trans>  
      <tr>Himmelsbläue</tr>  
    </trans>  
  </sense>  
</entry>
```

Weitere Beispiele

Beispiel 1:

```
(<POS: N.>) Winter<Gen.: m> (<Etym.: <impli.: aus d. Engl.><expli.: von engl <For.: winter>>>).
```

```
<entry id="9293702">  
  <etym lang="en" type="impl">aus d. Engl.</etym>  
  <etym lang="en" type="expl">von engl "winter"</etym>  
  <gramGrp>  
    <pos type="N"/>  
  </gramGrp>  
  <sense>  
    <trans>  
      <tr>Winter</tr>  
    </trans>  
  </sense>  
</entry>
```

Beispiel 2:

```
(<POS: N.>) Unterstreichung<Gen.: f> // {<Dom.: EDV>} Unterstrich<Gen.: m>; <SpecChar.: _> (<Expl.: engl.  
<For.: underscore>; <Expl.: ANSI-Code 0095>).
```

```

<entry id="49">
  <gramGrp>
    <pos type="N"/>
  </gramGrp>
  <sense>
    <trans lang="en">
      <tr>underscore</tr>
    </trans>
    <trans>
      <tr>Unterstreichung</tr>
      <def/>
    </trans>
    <trans dom="EDV">
      <usg type="hint">rel</usg>
      <tr>Unterstrich</tr>
    </trans>
    <trans>
      <usg type="hint">rel</usg>
      <tr>"_" (ANSI-Code 0095)</tr>
    </trans>
  </sense>
</entry>

```

Beispiel 3:

```

(<POS: N.>) IgNobel-Preis<Gen.: m> (<Def.: „<iron.: Nobelpreis> für die seltsamsten wissenschaftl. Leistungen>;
<Etym.: von engl. <For.: ignoble>"= „<Transl.: unehrenhaft>">).

```

```

<entry id="53">
  <etym lang="en" origin="ignoble">von engl. "ignoble"= „unehrenhaft"</etym>
  <gramGrp>
    <pos type="N"/>
  </gramGrp>
  <sense>
    <trans>
      <tr>IgNobel-Preis</tr>
      <def>"Nobelpreis" für die seltsamsten wissenschaftl. Leistungen</def>
    </trans>
  </sense>
</entry>

```

Beispiel 4:

```

<POS: N.> <Dom: Bot.> Blumenkohl<Gen.: m>; Karfiol<Gen.: m> (<Etym.: von engl. <For.: cauliflower>;
<Scientif.: Brassica oleracea var. botrytis>).

```

```
<entry id="4626535">
  <etym lang="en" origin="cauliflower">von engl. "cauliflower"</etym>
  <gramGrp>
    <pos type="N"/>
  </gramGrp>
  <usg type="dom">Bot</usg>
  <sense>
    <trans>
      <tr>Blumenkohl</tr>
      <def/>
    </trans>
    <trans lang="scientific">
      <tr>Brassica oleracea var. botrytis</tr>
    </trans>
    <trans>
      <tr>Karfiol</tr>
    </trans>
  </sense>
</entry>
```