

GENEALOGIES OF KNOWLEDGE

How to clean up documents and prepare them for the creation of an Arabic corpus

Purpose: To produce a machine-readable document that corresponds to the contents of the book, article or other item to be incorporated into the Genealogies of Knowledge Arabic corpus.

This is a two-stage process: first, you will need to 'clean' the document in Microsoft Word (i.e. ensure that the spelling and paragraphing corresponds to that of the original document); second, you will need to use an xml editor (Notepad++ or Oxygen, for example) to annotate the text with metadata. Notepad++ is used here for illustration.

STAGE ONE: CLEANING THE WORD FILE

1. If the text on which you are working is not already contained in a word document (.doc), you will need to convert it using Optical Character Recognition (OCR) software such as that offered by Google Docs:

Google Docs: <https://support.google.com/docs/thread/3593101?hl=en&msgid=4212188>

Please note that Google OCR needs as clean a pdf as possible. In other words, remove any borders, margins or fuzziness in the page.

2. Save this .doc file under the filename that you have been given for this document (e.g. mar000005). Please note that all filenames in the modern corpus should begin with a predefined prefix (such as 'mar' for the Modern Arabic corpus) and then a 6-digit number. These filenames will be allocated by the project manager.
3. You can delete any of the copyright and publication information that is often included on the inside front cover of books. This information will be included in the header file when we come to the second stage of the process. You should also delete the index and bibliography.
4. Make sure the text in the .doc file corresponds accurately with the contents of the original book. Make sure that spellcheck is activated and that you check all the underlined words/phrases. Always check with the original document, if in doubt.

If the .doc file has been created by scanning the original it will likely contain many character recognition (OCR) errors. See for instance, the example pasted below which has been OCR'd using Google Docs.

حكاية المرأة التي تلبس كسوة الريش

إذا ما أبصرتُموني ذات يوم صدفة في مطار الدار البيضاء أو على متن باخرة تغادر ميناء طنجة، ستعتقدون بأنني امرأة جد واثقة من نفسها. ولكنني لست كذلك البتة. حتى اليوم وفي هذا السن، أستشعر الرعب حين تراودني فكرة مغادرة المغرب واجتياز الحدود، لأنني أخشى تضییع الهدف الأسمى من السفر، أي فهم الأجانب الذين سأصادفهم في طريقي.

«إن السفر فرصة رائعة لترويض النفس ولاكتساب حكمة فريدة من نوعها، يكفي أن تركز انتباهك على ما يقوله الأجانب، وكلما فعلت ذلك تقوى نفوذك، فأحسن وسيلة تمنحك هذا النفوذ على أولئك الأجانب وغيرهم هي قدرتك على فهمهم». ذلك ما كانت تقوله جدتي الياسمين، وهي امرأة لا تعرف القراءة والكتابة، أمضت حياتها في حريم، أي في مجال أبوابه موصدة وآفاقه مسدودة. إنها أبواب كان يفترض أن لا

تفتحها النساء قط : «عليك أن تركز انتباهك على الأجنبي العابر وتحاولي فهمه. كلما فهمت أجنبيا عرفت ذاتك أكثر، وكلما عرفت ذاتك بشكل أفضل تقوت قدرتك على تحصينها، وازدادت حظوظك في الدفاع عن نفسك وإقناع الآخرين بقدراتك واكتساب جهم».

حكاية المرأة التي تلبس كسوة الريش

إن السفر فرصة رائعة لترويض النفس ولاكتساب حكمة فريدة من نوعها، يكفي أن تركزي انتباهك على ما يقوله الأجنبي، وكل ما فعلت ذلك تقوي نفوذك، فأحسن وسيلة تمنحك هذا النفوذ على أولئك الأجانب وغيرهم هي قدرتك على فهمهم. «ذلك ما كنت تقول جدي الياسمين، وهي امرأة لا تعرف القراءة والكتابة، أمضت حياتها في حريم، أي في مجال أبوابه موصدة وأفاقه مسدودة. إنها أبواب كان يفترض أن

العابرة المكسورة الجناح

تفتحها النساء قط : عليك أن تركزى انتباهك على الأجنبي العابر وتحاولي فهمه .لما فهمت أجنبيا عرفت ذاتك أكثر ، وكلما عرفت ذاتك بشكل أفضل تقوت ...

The overall look of the text can be improved by applying one font typeface and size to the whole text (Times New Roman for example).

5. Make sure that the paragraphs in the .doc file correspond to those in the document. Clicking on the 'Show ¶' button within Word's main toolbar will help you to identify more easily the end of each paragraph in the .doc file.
6. There should be no empty lines in the .doc file. For example, below is how the .doc file should look, even though in the book, there are several line breaks and footnotes in between.

منذ عدة سنوات، تجولت في عدد من المدن الغربية لتقديم كتابي «نساء على أجنحة الحلم» (*) الذي ترجم إلى ثمان وعشرين لغة. خلال تلك الجولة التي استجوبت فيها من طرف أكثر من مائة صحفي، لاحظت بأن الرجال يبتسمون كلما دعتهم الضرورة إلى النطق بكلمة «الحريم». صدمتني تلك الابتسامات. كيف

(*) صدرت الطبعة العربية من الكتاب عن: المركز الثقافي العربي ومنشورات الفنك.

7 حكاية المرأة التي تلبس كسوة الريش

يمكن أن نبتسم ونحن ننطق بكلمة ترادف كلمة السجن ؟ بل إنني وأكثر من ذلك كنت أستنكر الأمر .

منذ عدة سنوات، تجولت في عدد من المدن الغربية لتقديم كتابي «نساء على أجنحة الحلم الذي ترجم إلى ثمان وعشرين لغة. خلال تلك الجولة التي استجوبت فيها من طرف أكثر من مائة صحفي، لاحظت بأن الرجال يبتسمون كلما دعتهم الضرورة إلى النطق بكلمة «الحريم». صدمتني تلك الابتسامات. كيف يمكن أن نبتسم ونحن ننطق بكلمة ترادف كلمة السجن ؟ بل إنني وأكثر من ذلك كنت أستنكر الأمر .

7. If at any point there is a table or diagram in the text, you will need to re-arrange the formatting so that each line of the table or diagram exists as a separate paragraph. A table in the document that looks like this, for example:

الملحق : الجداول
جدول رقم (٢٢) تقديرات أعداد (بالآلاف) ونسب (%) الأميين بين السكان الشباب (١٥-٢٤ سنة) في الدول العربية حسب الجنس في الفترة بين ١٩٨٠ و ٢٠١٥ (١)

الدولة	١٩٨٠		١٩٩٠		٢٠٠٤-٢٠١٠		٢٠١٥	
	ذكور	إناث	ذكور	إناث	ذكور	إناث	المجموع بالآلاف	% إناث من المجموع
الأردن	١٢	٢٨	١٠	١٦	٥	١	٥	٣٧
عدد								
%	٤.١	١٤.٣	١.٩	٣.٦	٠.٧	٠.٢		

1

¹ From رقيقة سليم حمود by تعليم الإناث في الدول العربية 2008 p. 434

must be rendered like this:

الملحق : الجداول									
جدول رقم (٢٢) تقديرات أعداد (بالآلاف) ونسب (%) الأميين بين السكان الشباب									
(15-24 سنة) في الدول العربية حسب الجنس في الفترة بين 1980 و 2015 (1)									
الدولة 1980 1990 2000-2004 2015									
ذكور إناث ذكور إناث ذكور إناث المجموع بالآلاف % إناث من المجموع									
الأردن عدد 12 38 10 16 5 1 5									
27 0.2 0.7 3.6 1.9 14.3 4.1 %									

8. Return any text placed in columns to a standard, single-column format and remove any headers/footers from the document (e.g. page numbers, place markers etc.).
9. When you are finished, select all of the text (Ctrl + A) and go to Styles → drop-down menu → Clear All.
10. If the work contains any paratext (introductions, prefaces, endorsements, acknowledgements, forewords, afterwords, appendices, epigraphs, footnotes or endnotes), this should now be cut out from the main text and pasted into a separate Word document. These paratextual elements will be placed in separate sections of the XML file when we come to the xml annotation stage.

Introductions which essentially form chapter 1 of a work written by the original author should not be removed. Only truly paratextual elements – which stand apart from and in some sense seek to frame the main body of the text – should be cut and pasted into this new Word document.

11. Finally, save the main text document and the paratext document as plain text files (.txt) by going to FILE > SAVE AS.

You should use the same filename (e.g. mar000005) for the main text file; label the paratext file appropriately so that you can easily retrieve it later in the process (e.g. mar000005_paratext).

After you click Save, a dialogue box will open asking how you want to convert the document. Select 'Other Encoding' and choose 'Unicode (UTF-8)' from the menu on the right-hand side.

IMPORTANT: Make sure all of these files are in the same folder in your computer as the files 'gokheader.dtd' and 'goktext.dtd'. It is with these dtd files that the software will check for xml validation errors in the next stage of the process.

STAGE TWO: USING NOTEPAD++ TO ADD METADATA:

If you do not already have Notepad++ installed on your computer, it can be downloaded for free here:

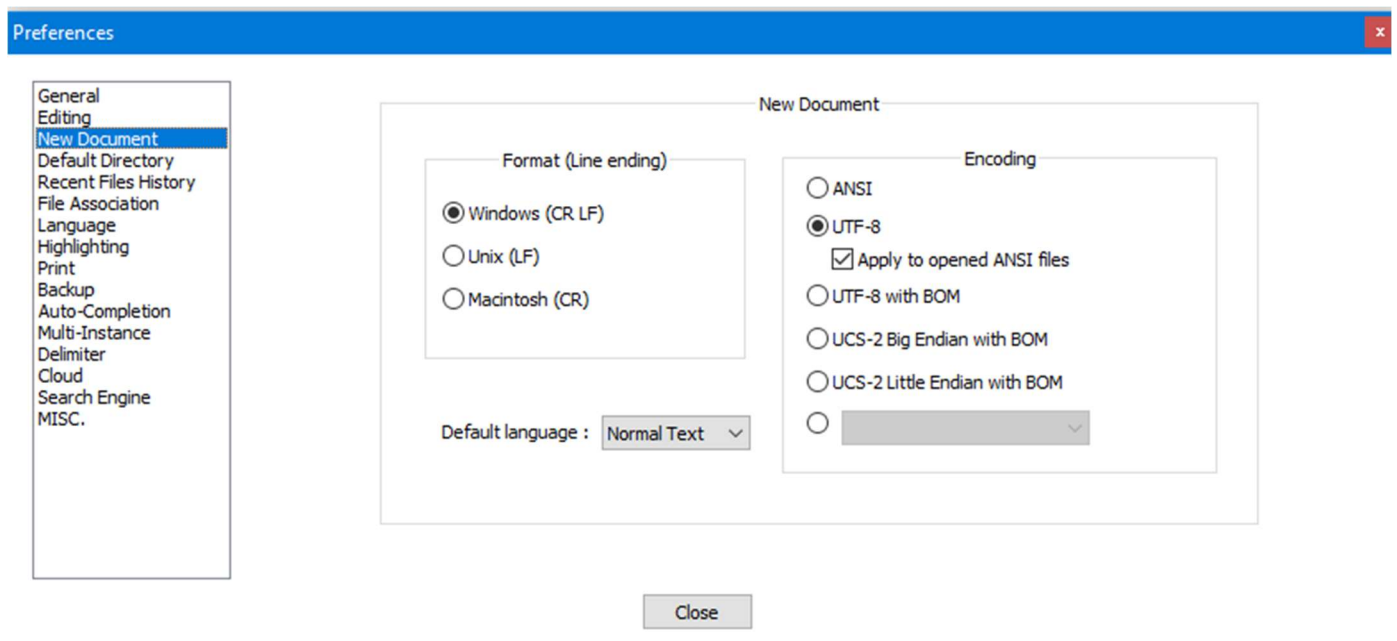
<https://notepad-plus-plus.org/downloads/>

Setting Up Notepad++

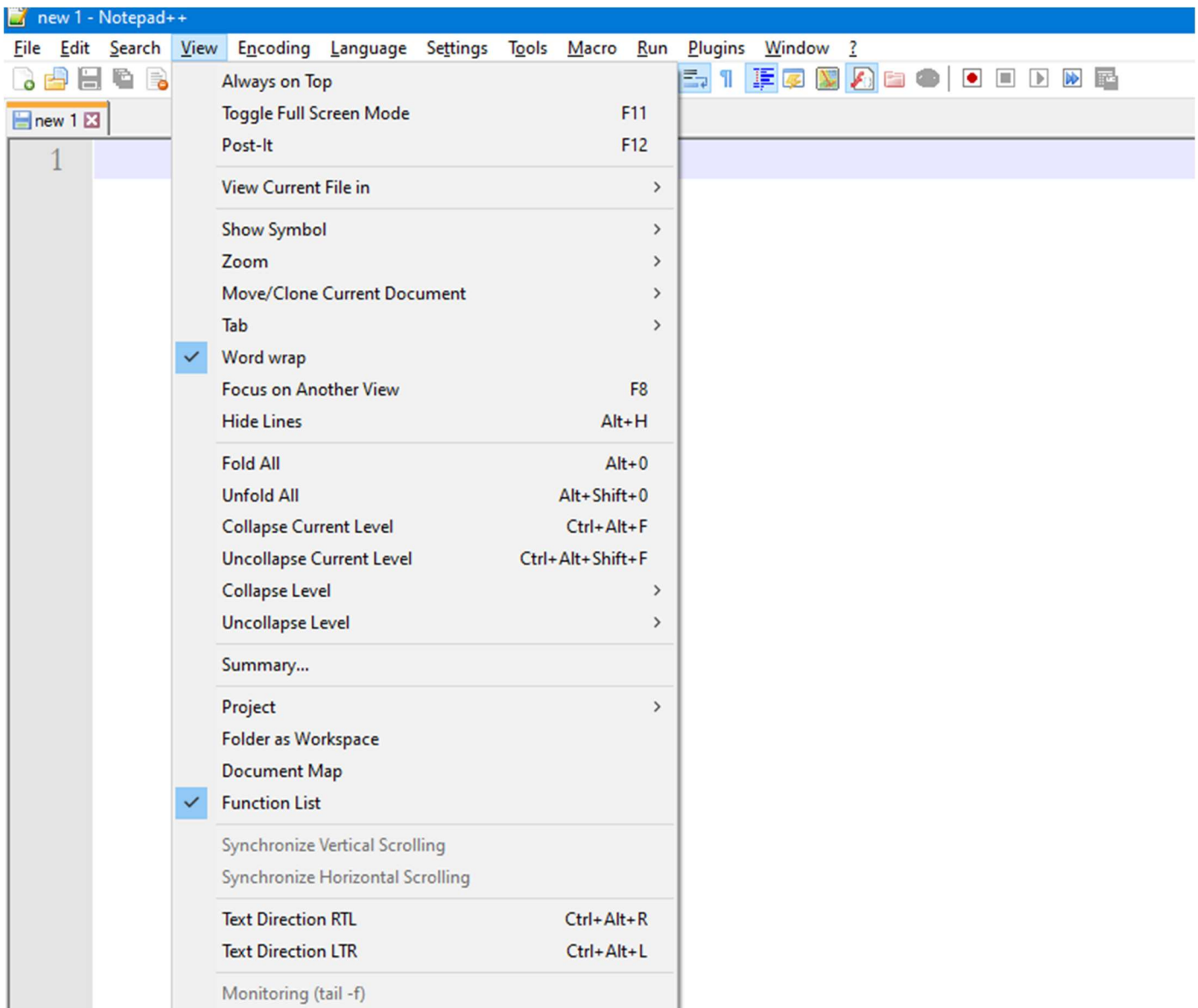
(NB you will only have to do this once, the first time you use Notepad++ on your machine.)

From the menu bar:

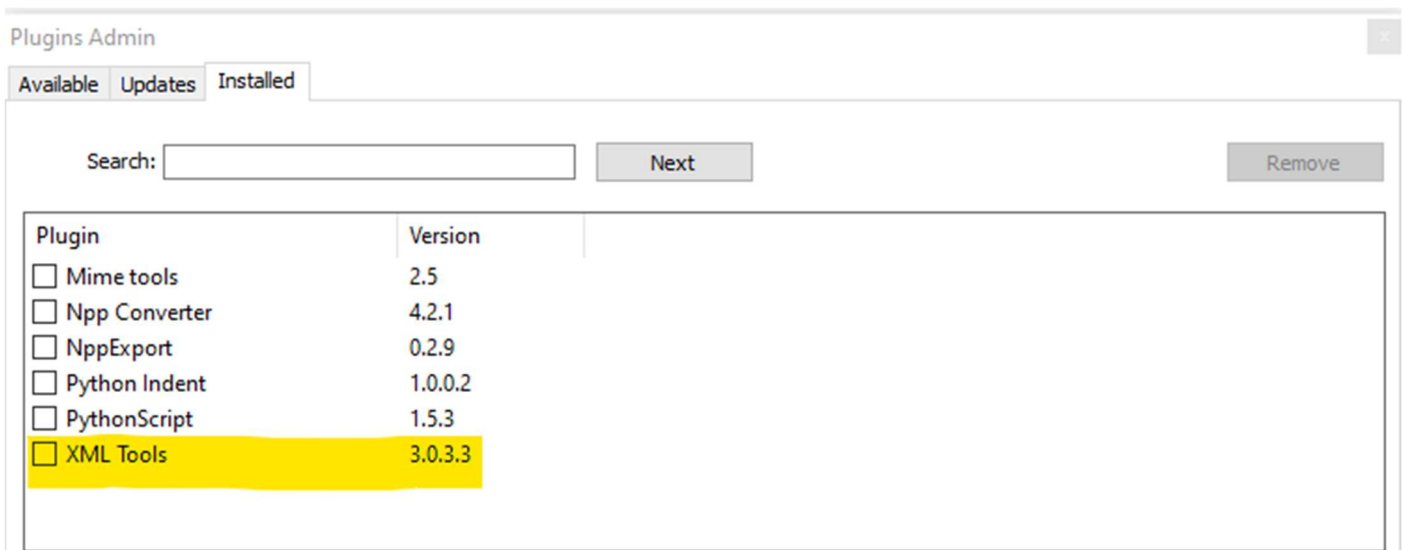
1. Select SETTINGS → Preferences...
2. Select NEW DOCUMENT (on the left-hand side)
3. Under ENCODING, make sure that you have UTF-8 selected, with the option to “APPLY TO OPENED ANSI FILES” checked



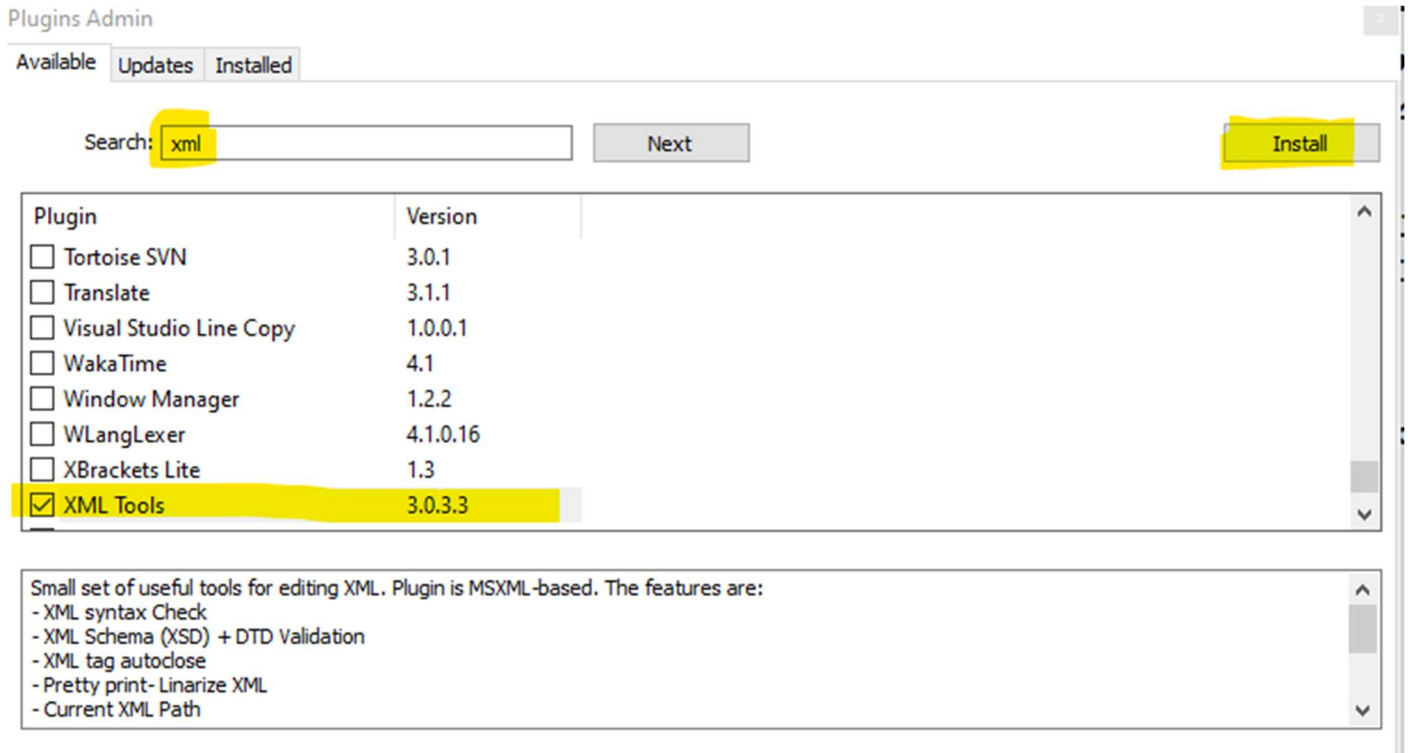
4. Select the View menu
5. Make sure that 'Word wrap' is selected.
6. Make sure that 'Function List' is selected.



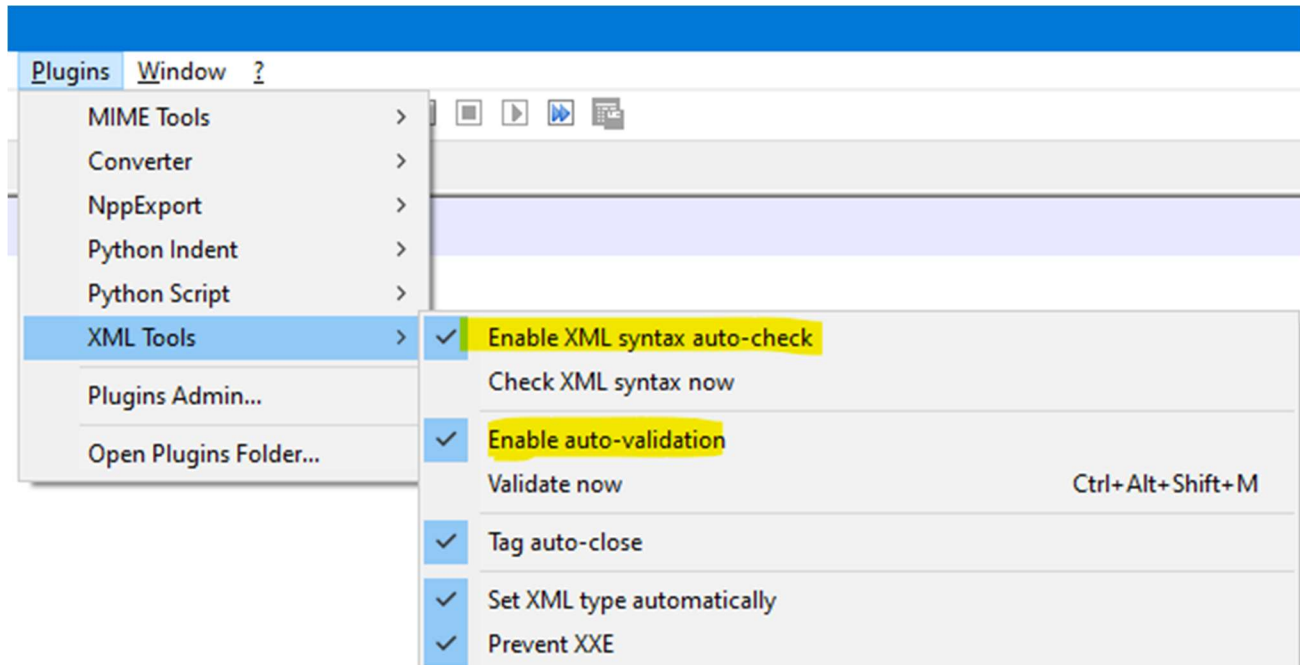
7. Select the Plugins menu, and click on Pugins Admin
8. Select the 'Installed' tab and check that the plugin 'XML Tools' is installed.



9. If the plugin is not installed, go to the 'Available' tab. Type xml and click Next. Tick the box next to 'XML tools' then click 'Install'. Notepad++ will ask for you confirmation then it will restart.



10. Once Notepad++ restarts, go to 'Plugins' menu, select 'XML Tools' and tick the following two options: 'Enable XML syntax auto-check' and 'Enable auto-validation'.



It is often useful to have a printed copy of the 'goktext.dtd' and 'gokheader.dtd' files to hand when adding the metadata. These documents include all the possible metadata tags that you can use to annotate the text and it is against these .dtds that Notepad++ will check for errors.

Creating the xml text file:

11. Open your main .txt file in Notepad++ by going to FILE > OPEN. Next, go to FILE > SAVE AS and save it as an .xml file (e.g. mar000005.xml).

12. Insert the following lines at the top of this file:

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!DOCTYPE goktext SYSTEM "goktext.dtd">
<goktext>
<section id='s1'>
```

At the very end of the text, place these lines:

```
</section>
</goktext>
```

13. Place the title between <title>...</title> tags and any subtitles and/or subheadings between <shead>...</shead> tags.

Important: All metadata tags need to have two parts, one that opens them and one that closes them. A forward slash / is used to distinguish a closing tag from an opening tag (e.g. <title>.....</title>).

14. Each chapter should be placed between <chapter>...</chapter> tags and numbered in the correct order, as follows:

```
<chapter n="1">....</chapter>
<chapter n="2">....</chapter>
<chapter n="3">....</chapter>
```

...

15. If you need to edit the Arabic text in any way, you need to switch the view to become Right-to-Left: go to 'View' → and choose 'Text Direction RTL' in order to view Arabic texts properly. Be careful as this will also switch the display of the metadata to become right-to-left as well. Use this feature only to work on the Arabic text itself. If you want to work with tags, switch back to the original viewing option (Go to 'View' → and choose 'Text Direction LTR').

16. Please note that the tag <introduction> should NOT be used to refer to the first chapter of a book. This should only be used in cases where the introduction has been written by e.g. the translator or editor.

17. Go through the text and place the tag <p/> at the end of every paragraph in the book. Please see the instructions on using Regular Expressions for a quick and easy means of doing this:

<http://genealogiesofknowledge.net/genealogies-knowledge-corpus/corpus-text-preparation/>

The easiest way to insert the tag <p/> at the end of every paragraph is to highlight the text you want to apply changes to, then press Ctrl+H, then in the 'Find what' box type \r\n and in the 'Replace with' box type <p/>\r\n.

Note: <p/> is a special case in that, unlike all other tags, it does not require an opening tag at the beginning of the paragraph. Simply place <p/> at the end of each paragraph.

18. Any tables must be placed between <list type="tables">...</list> tags.

19. Text appearing in any diagrams can be placed between <diagram>...</diagram> tags.

20. Any web addresses (URLs) should be placed between <url>...</url> tags. Please see the instructions on using Regular Expressions for a quick and easy means of doing this:

<http://genealogiesofknowledge.net/genealogies-knowledge-corpus/corpus-text-preparation/>

Adding in the paratextual elements

Once the main body of the text has been annotated with XML tags, we can now add the paratextual elements (translator's preface, footnotes, etc.) in a series of further sections within the XML file.

21. Section s1 should now contain the main body of the text, excluding any paratextual elements such as introductions, prefaces, footnotes etc..
22. At the end of this section (after the </section> tag), you will now need to create a new section <section id='s2'>.
23. In this new section, you need first to insert the following tag: <frontmatter id='p1'>. Then you can copy and paste any frontmatter contained within the volume (introductions, prefaces, endorsements, acknowledgements, forewords) into this new section.
24. Make sure each of these is tagged with the corresponding metadata element. For example, introductions should be tagged with <introduction>...</introduction>.

```
22 | </section>
23 | <section id='s2'>
24 | <frontmatter id='p1'>
25 | <introduction>
26 | <shead>مقدمة</shead>
27 | نص المقدمة<p/>
28 | </introduction>
29 | </frontmatter>
30 | </section>
```

25. Close section s2 with a </frontmatter> and a </section> tag and then create a new section <section id='s3'>.
26. In this section, you must first add the following: <footnote id='p2'>.
27. You can then copy and paste any footnotes that you extracted from the text into this new section. Note that these do not need to be tagged individually. You should simply add a <p/> tag after each.

```
30 | </section>
31 | <section id='s3'>
32 | <footnote id='p2'>
33 | <p/> صدرت الطبعة العربية من الكتاب عن: المركز الثقافي العربي ومنشورات الفنك (2)
34 | </footnote>
35 | </section>
```

28. Close section s3 with a </footnote> tag and a </section> tag and then create a final section <section id='s4'>.
29. In this section, you must first add the following: <backmatter id='p3'>.
30. Here, you can copy and paste any backmatter contained within the volume (endnotes, afterwords, appendices, epigraphs).
31. Make sure each of these is tagged with the corresponding metadata element. For example, endnotes should be tagged with <endnote>...</endnote>.

32. Close section s4 with a `</backmatter>` tag. The `</section>` tag should already be in place so no need to add it again here.
33. Finally, we need to add 'anchors' into section s1 with the corresponding id numbers for each of these paratextual elements. `<anchor id="p1"/>` (corresponding to the frontmatter) should be added to the very beginning of section s1. `<anchor id="p2"/>` (corresponding to the footnotes) and `<anchor id="p3"/>` (corresponding to the backmatter) can be added to the very end of section s1.
34. To check for errors in the metadata, click on Plugins → XML Tools → Check XML syntax now, to see them. The most common errors are when the closing tag (e.g. `</thead>`) has been omitted or if you have used/misspelled a tag/element that is not included in the `goktext.dtd`.

Creating the header file

For the last step of the process, we need to create a header file (.hed) for the corpus text. This provides all the metadata relevant to this specific text: the date of publication, the names of the author(s), translator(s), etc.. Including this information helps the corpus analyst to refine their corpus selection according to more specific criteria. The header file needs to reflect accurately the structure of the xml text file. It will be used to help the software to distinguish the main body of the text (contained in section s1) from the paratext (contained in sections s2, s3 and s4).

35. To begin, click FILE > NEW and then save this new file as a .hed file with the same filename as the text file (e.g. `mar000005.hed`).
36. Paste the following lines into this file:

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!DOCTYPE gokheader SYSTEM "gokheader.dtd">
<gokheader>
  <document format="" subcorpusid="Modern_Arabic" filename="">
    <collection_title></collection_title>
    <summary></summary>
  </document>
  <section id="s1" publication_date="" authorship_date="" outlet=""
    language="">
    <title></title>
    <author>
      <name></name>
    </author>
    <translation status="translation" language="English">
      <translator certainty="certain">
        <name></name>
      </translator>
      <source date="" language=""><original_title></original_title></source>
    </translation>
    <copyright status="">
      <terms></terms>
      <holder></holder>
    </copyright>
    <comments></comments>
  </section>
</gokheader>
```

37. Insert relevant information in each field (e.g. author, title, filename, summary etc.).

In cases where you want to add more than one name (e.g. translators) just copy and paste that line/section. For example, a book with one translator would look like this:

```
<translator>
  <name>فاطمة الزهراء أزرويل</name>
</translator>
```

Whereas a book with two translators will look like this:

```
<translator>
  <name>فاطمة الزهراء أزرويل</name>
  <name>المترجم الثاني</name>
</translator>
```

38. Please note that for most texts included in corpus, the authorship_date of the translation and the publication_date of the translation will be the same. In some cases, however, the publication_date will be different if, for example, it was re-issued by another publisher at a later date.

39. Arabic book titles should be kept in Arabic within the <title></title> tag, but a transliteration of the Arabic title should be included in parantheses within the same element, e.g.

```
<title>شهرزاد ترحل إلى الغرب (Shahrazād Tarḥalu ilā al-gharb)</title>.
```

For transliteration, we use the Library of Congress system explained in full at their website

<https://www.loc.gov/catdir/cpsd/romanization/arabic.pdf>.

40. For translations, the <original_title> element should be filled with the original title transliterated in the roman alphabet e.g. for the Arabic translation of *Scheherazade Goes West* the Arabic and English titles should be tagged as follows

```
<title>شهرزاد ترحل إلى الغرب</title>
```

```
<original_title>Scheherazade Goes West</original_title>
```

Here's an example of what a header file should look like when this first step has been completed:

```

1  <?xml version="1.0" encoding="UTF-8" standalone="no"?>
2  <!DOCTYPE gokheader SYSTEM "gokheader.dtd">
3  <gokheader>
4  <document format="book" subcorpusid="Modern_Arabic" filename="ara000005">
5    <collection_title>شهرزاد ترحل إلى الغرب</collection_title>
6    <summary>يتناول الكتاب مفهوم المرأة عبر شخصية شهرزاد وكيف ظلت كتاباً مفتوحاً يختلف الناس فيه فتارة هي حبيسة القصور وتارة فارسة</summary>
7  </document>
8  <section id="s1" publication_date="2008" authorship_date="2008" outlet="publisher"
9    language="Modern_Arabic">
10    <title>شهرزاد ترحل إلى الغرب (Shahrazād Tarḥalu ilā al-gharb)</title>
11    <author>
12      <name>فاطمة المريني</name>
13    </author>
14    <translation status="translation" language="Modern_Arabic">
15      <translator certainty="certain">
16        <name>فاطمة الزهراء أزرويل</name>
17        <name>المترجم الثاني</name>
18      </translator>
19      <source date="2008" language="English"><original_title>Scheherazade Goes West</original_title></source>
20    </translation>
21    <copyright status="copyright_granted">
22      <terms>Not available for commercial purposes</terms>
23      <holder>المركز الثقافي العربي</holder>
24    </copyright>
25    <comments></comments>
26  </section>
27  <section id="s2" publication_date="2008" authorship_date="2008" outlet="publisher"
28    language="Modern_Arabic">
29    <title>Frontmatter</title>
30    <author>
31      <name>فاطمة المريني</name>

```

41. If the text you are working with is not a translation (i.e. a document originally written in Arabic), you will need to delete the whole section of the header relating to translation from the header (<translation>...</translation>), to ensure that the header file validates against the gokheader.dtd.
42. Now we need to create a new section in the header file to correspond with the structure of the xml text file. After the </section> tag in the header file, copy and paste the following:

```

<section id="s2" language="" outlet="" authorship_date="" publication_date="">
  <title>Frontmatter</title>
  <author><name></name></author>
  <copyright status="">
  <terms></terms>
  <holder></holder>
</copyright>
<comments></comments>
</section>
<section id="s3" language="" outlet="" authorship_date="" publication_date="">
  <title>Footnotes</title>
  <author><name></name></author>
  <copyright status="">
  <terms></terms>
  <holder></holder>
</copyright>
<comments></comments>
</section>
<section id="s4" language="" outlet="" authorship_date="" publication_date="">
  <title>Backmatter</title>
  <author><name></name></author>
  <copyright status="">

```

```

    <terms></terms>
    <holder></holder>
</copyright>
<comments></comments>
</section>

```

43. Insert the relevant information in each field for each section. For example, if the introduction was originally written by the translator, his/her name should appear in the author name field. Some of the metadata will be the same for every section (e.g. the copyright information) but others will be different (e.g. title, author, authorship date etc.).

44. Finally, check for errors in the same way as above for the xml text file: click on Plugins → XML Tools → Check XML syntax now.

Dealing with edited collections

Because edited collections (unlike monographs) generally contain many different texts written at different times and/or by different authors, we cannot place all of the volume's content in a single section. Instead, each individual text contained in the collection must be placed in its own section within the xml text file, as follows:

```

<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!DOCTYPE goktext SYSTEM "goktext.dtd">
<goktext>
    <section id="s1">
        [Insert first text here]
    </section>
    <section id="s2">
        [Insert second text here]
    </section>
    <section id="s3">
        [Insert third text here]
    </section>
    ...
</goktext>

```

Important: This will now mean that the section numbers for the frontmatter, footnotes and backmatter will no longer be s2, s3 and s4. Instead you will have to paste in these sections after all of the main body of the volume has been added. If for example the main body of the text contains four texts written by different authors at different times, these will be labelled s1, s2, s3 and s4. The frontmatter will then be s5, the footnotes s6 and the backmatter s7.

Make sure the structure of the header file accurately reflects the structure of this xml file.

Validation

Selecting the option to 'Enable auto-validation' (see point 10 above) means that Notepad++ will give an error message if there is something wrong with the annotation of the file. The error messages are self-explanatory and often refer to missing closing tags (e.g. </chapter>) or the presence of a tag or a value not accepted by the dtd files (e.g. using "Arabic" instead of "Modern_Arabic" in the tag <section id="s1" publication_date="2008" authorship_date="2008" outlet="publisher" language="Modern_Arabic">). Once your file is error-free, Notepad++ may display a message asking for an XML schema (XCD) as below. Please ignore this message and click cancel as it does not affect the validation of the file.

Example header file:

```
2 <!DOCTYPE gokheader SYSTEM "gokheader.dtd">
3 <gokheader>
4 <document format="book" subcorpusid="Modern_Arabic" filename="ara000005">
5 <collection_title>شهرزاد ترحل إلى الغرب</collection_title>
6 <summary>يتناول الكتاب مفهوم المرأة عبر شخصية شهرزاد وكيف ظلت كتابا مفتوحا يختلف الناس فيه فتارة هي حبيسة القصور وتارة فارس</summary>
7 </document>
8 <section id="s1" publication_date="2008" authorship_date="2008" outlet="publisher"
9 language="Modern_Arabic">
10 <title>شهرزاد ترحل إلى الغرب (Shahrazād Tarḥalu ilā al-gharb)</title>
11 <author>
12 <name>فاطمة المريني</name>
13 </author>
14 <translation status="translation" language="Modern_Arabic">
15 <translator certainty="certain">
16 <name>فاطمة الزهراء أزرويل</name>
17 <name>المترجم الثاني</name>
18 </translator>
19 <source date="2008" language="English"><original_title>Scheherazade Goes West</original_title></source>
20 </translation>
21 <copyright status="copyright_granted">
22 <terms>Not available for commercial purposes</terms>
23 <holder>المركز الثقافي العربي</holder>
24 </copyright>
25 <comments></comments>
26 </section>
27 <section id="s2" publication_date="2008" authorship_date="2008" outlet="publisher"
28 language="Modern_Arabic">
29 <title>Frontmatter</title>
30 <author>
31 <name>فاطمة المريني</name>
32 </author>
33 <translation status="translation" language="English">
34 <translator certainty="certain">
35 <name>فاطمة الزهراء أزرويل</name>
36 <name>المترجم الثاني</name>
37 </translator>
38 <source date="2008" language="English"><original_title>Scheherazade Goes West</original_title></source>
39 </translation>
40 <copyright status="copyright_granted">
41 <terms>Not available for commercial purposes</terms>
42 <holder>المركز الثقافي العربي</holder>
43 </copyright>
44 <comments></comments>
45 </section>
46 <section id="s3" publication_date="2008" authorship_date="2008" outlet="publisher"
47 language="Modern_Arabic">
48 <title>Footnotes</title>
49 <author>
50 <name>فاطمة المريني</name>
51 </author>
52 <copyright status="copyright_granted">
53 <terms>Not available for commercial purposes</terms>
54 <holder>المركز الثقافي العربي</holder>
55 </copyright>
56 <comments></comments>
57 </section>
58 </gokheader>
```