



Export-AddIn
Version 1.0
10.7.2012

Marc Reznicek
Humboldt-Universität zu Berlin



Content

1 Installation 3

2 Deinstallation..... 5

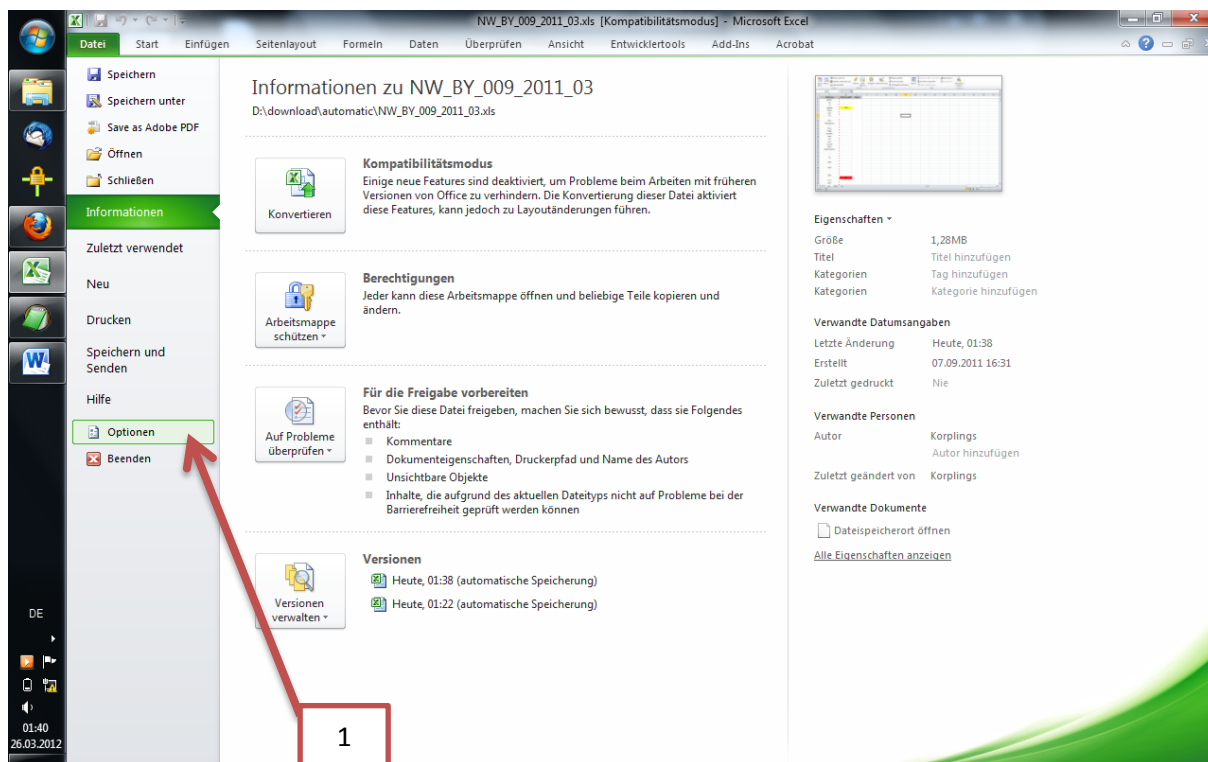
3 Create CONLL Sheet and export 6

4 Contact 10

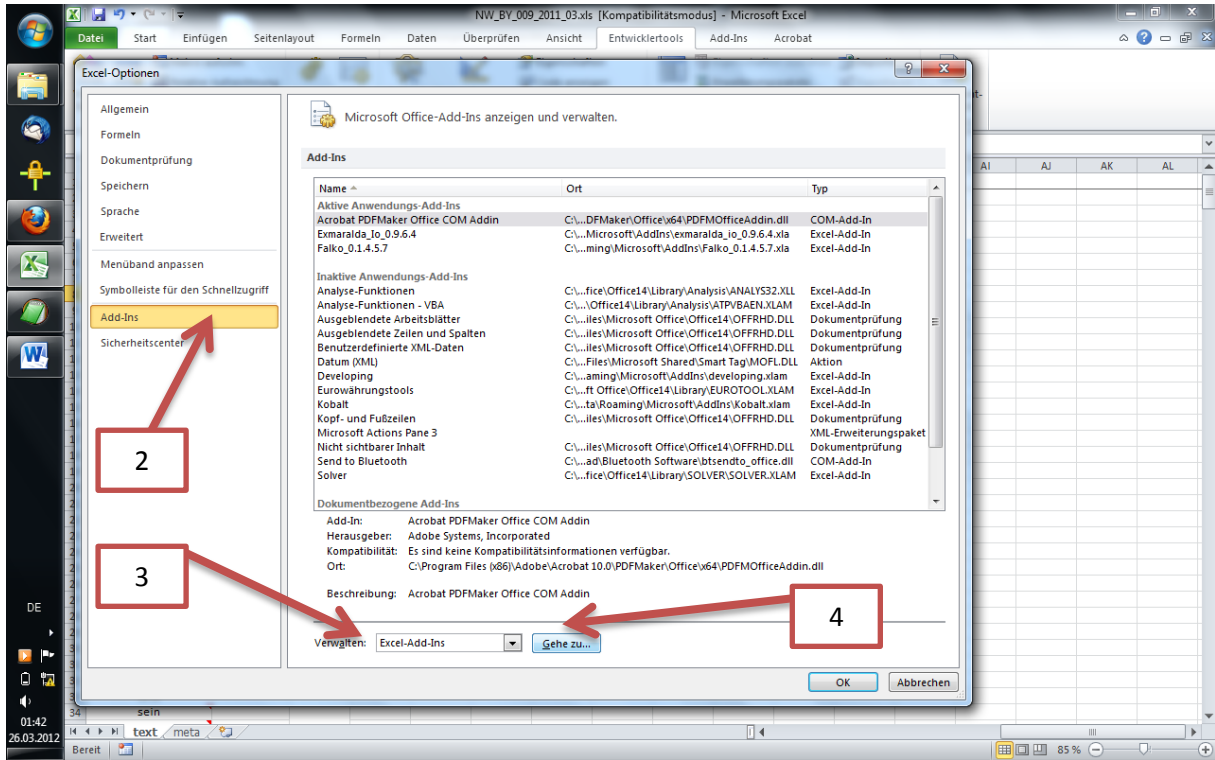
1 Installation

This installation demonstration is adapted to Microsoft Office 2010. Installation steps for older versions are not identical but very similar.

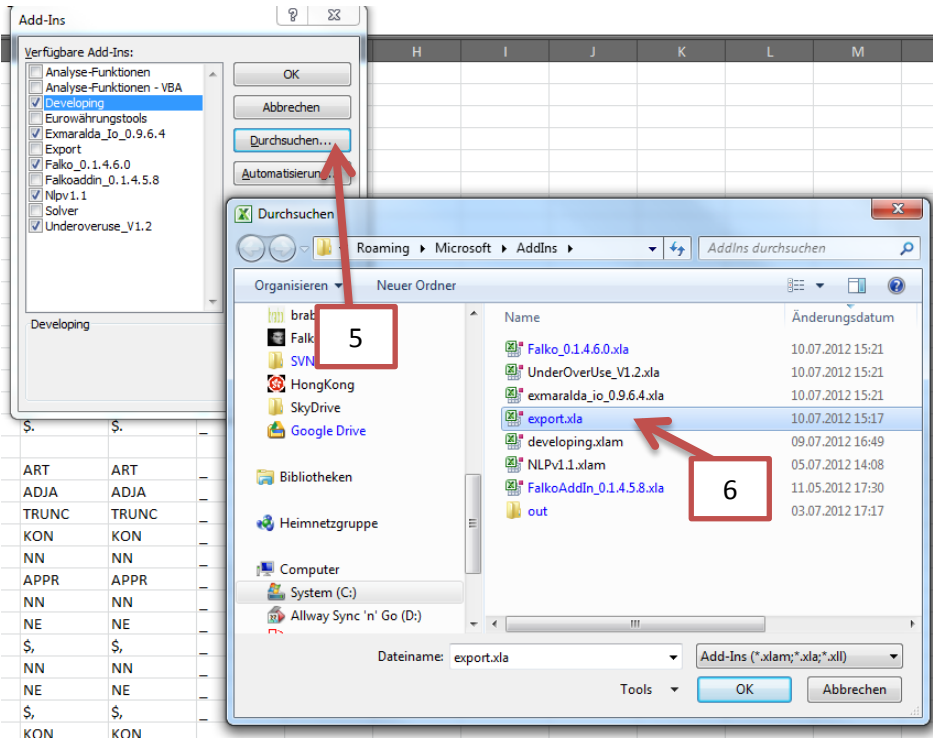
- 1) Open "File" Tab
- 2) Click "Options"



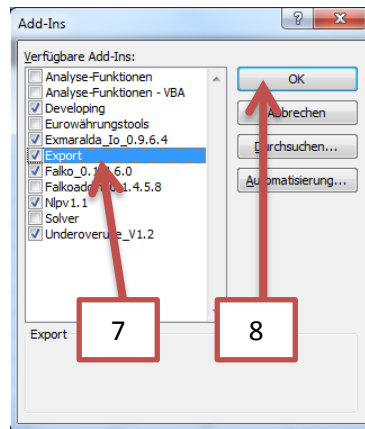
- 3) Choose the entry „AddIns“ in the left pane.
- 4) Choose "Excel-AddIns" in the drop-down box
- 5) Click "go to"



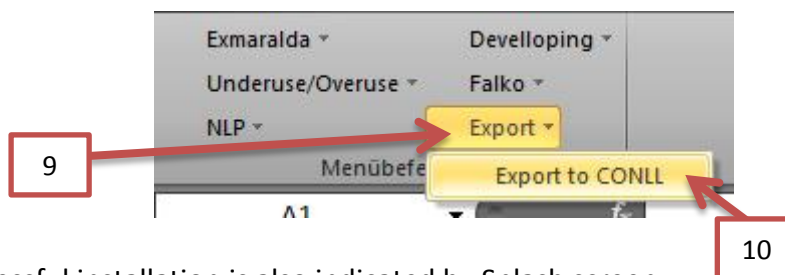
- 6) In the new formular click „search for“
- 7) Copy the export.xla into the opened folder
- 8) Choose export.xla and click “OK”



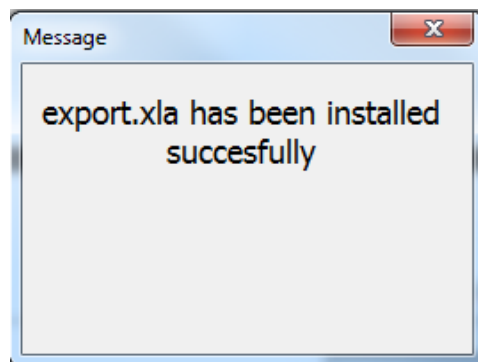
9) Mark export item and click “OK”.



10) If the Add-Ins Tab in Ribbon shows „Export“ with item „Export to CONLL“ the AddIn was installed correctly.

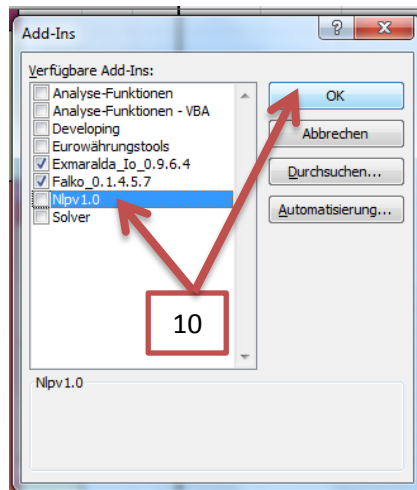


11) Successful installation is also indicated by Splash screen



2 Deinstallation

To remove the AddIn check installation step 5.



3 Create CONLL Sheet and export

The Add-In allows the export of data organized in the following way:

- a) A text layer is divided into one token per excel cell
- b) Each column contains a single layer.
- c) The first row contains the layer names
- d) Spans over more than one token inside a layer are dealt with correctly.

	A	B	C	D	E	F	G
1	dipl	clean	norm	pos	pos_cor	lemma	lemma_cor
2	I.	I.	I.	NE	NE	I.	I.
3	Der	Der	Der	ART	ART	d	d
4	weiffe	weisse	weiße	ADJA	ADJA	weiß	weiß
5	Thannen»						
6	baum	Thannenbaum	Tannenbaum	NN	NN	Tannenbaum	Tannenbaum
7	.	.	.	\$.	\$.	.	.
8	Abies	Abies	Abies	NE	NE	<unknown>	<unknown>
9	.	.	.	\$.	\$.	.	.
10	Namen	Namen	Namen	NN	NN	Name	Name
11	.	.	.	\$.	\$.	.	.
12	Figur	Figur	Figur	NN	NN	Figur	Figur

There are different CONLL-formats (<http://ilk.uvt.nl/conll/>) containing a different amount of layers (up to 14). Therefore the AddIn allows for the user to choose which layers should be exported and in which order.

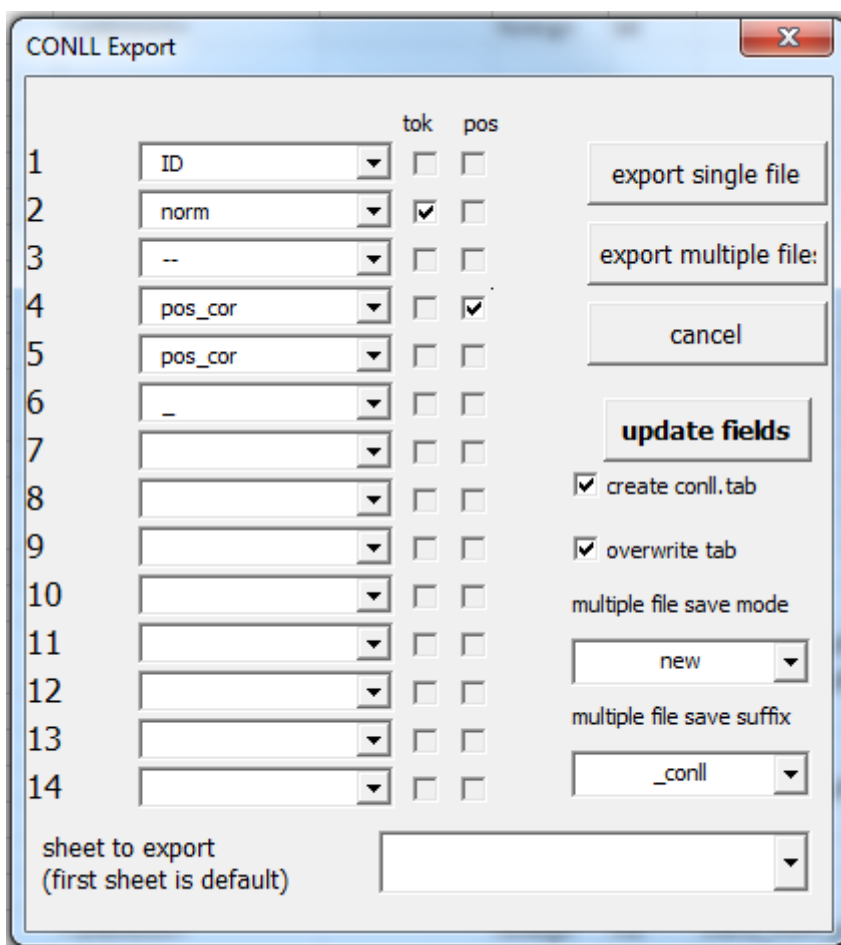
A 6-layer version is shown below.

ID	tok	lemma	POS1	POS2	DEP
1	I.	--	NE	NE	--
2	Der	--	ART	ART	--
3	weiße	--	ADJA	ADJA	--
4	Tannenbaum	--	NN	NN	--
5	.	--	\$.	\$.	--
1	Abies	--	NE	NE	--

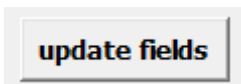


2 . -- \$. \$. _

- The IDs are repeated for every sentence and sentences are divided by empty lines.
- Once started a form allows the user to choose the different layers present in an open workbook. Layers may be ordered as you like and they may occur more than ones (as POS-tags are often repeated).
- The item “ID” includes the token counter in each sentence.
- To calculate the right sentence splits, the token and the POS level must be indicated by the checkbox to the right. There may be only one token layer and one POS layer



The dropdown boxes are automatically filled for the active worksheet. You may switch worksheets and update the fields and boxes by pressing the button “update fields”.

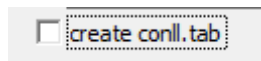


- The AddIn works in a two step manner
- 1) A CONLL-Spreadsheet is created

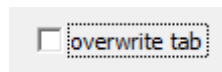
	A	B	C	D	E	F
1	1	I.	--	NE	NE	
2	2	Der	--	ART	ART	
3	3	weiße	--	ADJA	ADJA	
4	4	Tannenbaum	--	NN	NN	
5	5	.	--	\$.	\$.	
6						
37	19	an	--	APPR	APPR	
38	20	der	--	ART	ART	
39	21	Länge	--	NN	NN	

1) The CONLL-data might be exported

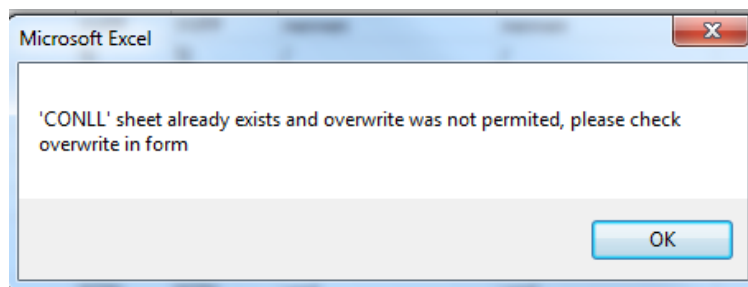
Unchecking the “create conll.tab” checkbox will skip the export into a conll.tab file.



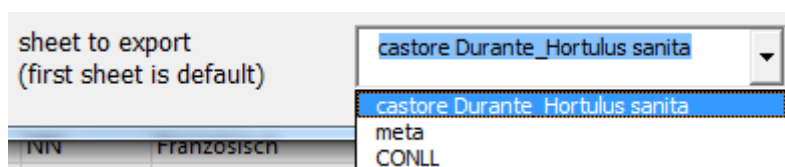
If the file already includes a “CONLL”-Worksheet, it will be deleted and replaced by default. If you do not wish that, uncheck the “overwrite tab” checkbox.



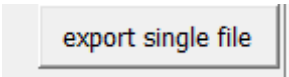
In case a “CONLL” sheet exists a warning will be thrown.



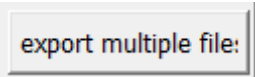
By default, data in the first worksheet of the active workbook will be exported. If you wish to change that, you might choose the name of the worksheet to be exported instead.



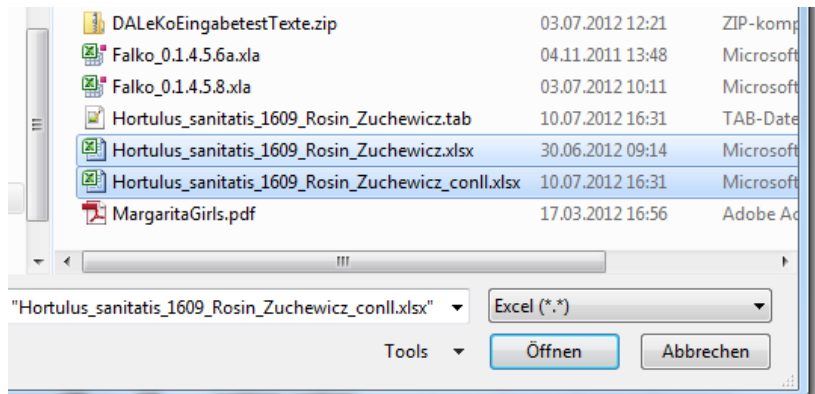
To export only the open workbook press “export single file”.



To export multiple files press “export multiple files” .

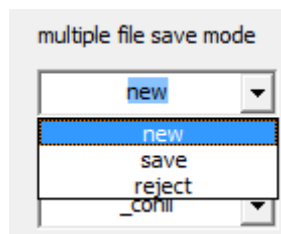


In multiple file export mode choose the files to be exported in the opening dialog.

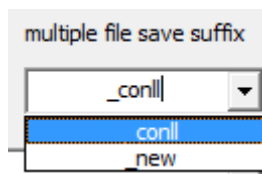


You may choose, if you wish to save the changes in each workbook:

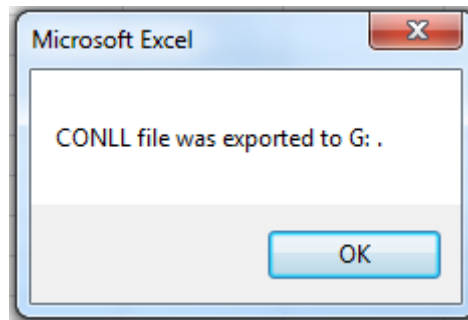
- “new” will save changes in a new workbook
- “save” will save changes in the old workbook
- “reject” will not save changes



If you save the changes in a new file you may choose a suffix under “multiple file save suffix” or you may type your own. (default is “_conll”)



Once all files are exported you get a message indicating the export folder. In the current version that’s allways the folder of the original data.



4 Contact

For questions, problems and request please contact:

Marc Reznicek

Marc.Reznicek@staff.hu-berlin.de