Multilingual features in SkE Terminology, EUR-Lex

Vít Baisa



vit.baisa@sketchengine.co.uk

7th Sketch Engine Workshop Portorož, May 23, 2016

Parallel concordance

OPUS2 English

configuration entry value from the given key . Command-line Options Amarok supports. Amarok-Konfigurationseintrag des angegebenen Schlüssels zurück . Optionen auf der Befehlszeile features of ssh , including public key authentication and compression . See such as passwords , cryptographic keys , your addressbook , documents private files define the same configuration key , the file with the highest precedence determines which value is used for the key . For example , if the following two groups . The first group contains the keys LargeCursor and SingleClick , the SingleClick . the second group contains the keys Show hidden files and Sort by : [Name Entries in a group consist of a key and value separated by an equals sign value separated by an equals sign . The key can contain spaces and may be followed on a key-by-key basis. If the same key within a certain group is defined defined in more than one place , the key value read from the directory tree have the highest precedence . If a key in a certain group is defined multiple locked down by adding [\$i] behind the key .& eg :: Color [\$i] = blue A expansion the value of a configuration key can be constructed from the value expansion for a configuration entry , the key must be followed by [\$e]. Normally the desktop is used to look up the key value . If the default language (orresponds to the selected language , the key entry without index is used . In the described along with the corresponding key GreetString in & kdmrc ; . You can specified by FifoDir = . The subdir is the key to addressing and security ; the sockets escaped contents for . dmrc . Unlisted keys will default to previously saved values

OPUS2 German

		und Datenkompression . Siehe auch
Beispiel Passwörter , kryptografische	Schüssel	, Ihr Adressbuch , Firmen interne
. Wenn in zwei Dateien der gleiche	Schüssel	definiert ist, wird der Wert des
. Wenn in zwei Dateien der gleiche	Schüssel	definiert ist , wird der Wert des
Gruppen . Die erste Gruppe enthält die	Schüssel	LargeCursor und SingleClick , die
Gruppen . Die erste Gruppe enthält die	Schüssel	LargeCursor und SingleClick , die
		/Wert-Paare getrennt von einem Gleichheitszeis
		/Wert-Paare getrennt von einem Gleichheitszeis
die Inhalte aller gefundenen Dateien	Schüssel	für Schlüssel kombiniert. Falls sich
die Inhalte aller gefundenen Dateien	Schüssel	für Schlüssel kombiniert. Falls sich
die Inhalte aller gefundenen Dateien	Schüssel	für Schlüssel kombiniert. Falls sich
werden , indem man [\$ i] hinter den	Schüssel	einfügt . Beispiel : Color [\$ i]
nutzen zu können , muss der jeweilige	Schüssel	mit dem Parameter [\$ e] versehen
nutzen zu können , muss der jeweilige	Schüssel	mit dem Parameter [\$ e] versehen
		können mit einem Sprachcode indexiert
\$ { USER } @ \$ { HOST } Die meisten	Schüssel	können mit einem Sprachcode indexiert
die zusammen mit dem dazugehörigen	Schüssel	GreetString in & kdmrc ; beschrieben
festgelegt wurde . Der Unterordner ist der	Schüssel	zur Adressierung und die Sicherheit
Inhalt für . dmrc . Nicht aufgeführte	Schüssel	sind mit den zuletzt gespeicherten

OPUS2 French

san, comme i l'autheritication par cel publique ou le compactage. Vui le semitibles, comme le moto de passe a les cité de chillinge, voi voi carriel d'adresses Quard deux fichiers définiser la même cel de configuration, le fichier apart Quard deux fichiers définiser la même cel de configuration, le fichier apart quard deux fichiers définiser la même cel de configuration, le fichier apart proges. Le premier groupe content les dels LageCursor et EngleCito, tandat d'un groupe es composant d'une cel et el d'une avait apartera par un	
Quand deux fichiers definissent la mème ciè de configuration, le fichier ayant Quand deux fichiers définissent la mème ciè de configuration, le fichier ayant groupes. Le premier groupe content les clès LangeCursor et SingleCitick, tandis groupes. Le premier groupe content les clès LangeCursor et SingleCitick, tandis	
Quand deux fichiers définissent la même clé de configuration, le fichier ayant groupes. Le premier groupe content les clés LargeCursor et SingleClick, tandis groupes. Le premier groupe content les clés LargeCursor et SingleClick, tandis	
groupes. Le premier groupe contient les clés LargeCursor et SingleClick, tandis groupes. Le premier groupe contient les clés LargeCursor et SingleClick, tandis	
groupes . Le premier groupe contient les clés LargeCursor et SingleClick , tandis	
d'un groupe se composent d'une clé et d'une valeur séparées par un	
d'un groupe se composent d'une clé et d'une valeur séparées par un	
onfiguration sont combinées selon un principe clé par clé . Si la même clé au sein d	
onfiguration sont combinées selon un principe clé par clé . Si la même clé au sein d	
onfiguration sont combinées selon un principe clé par clé . Si la même clé au sein d	
verrouillé en ajoutant [\$ i] à côté de la ché , & pex ; & # 160 ; : Color [\$ i	
expansion du shell , la valeur d'une clé de configuration peut être construite	
expansion du shell , la valeur d'une clé de configuration peut être construite	
bureau sert à rechercher la valeur de la clé . Si l ' on a choisi la langue par	
bureau sert à rechercher la valeur de la clé . Si l ' on a choisi la langue par	
sont décrites au même endroit que la clé correspondante (GreetString) dans	
FifoDir = . Le sous-dossier est la clé de l'adressage et de la sécurité	
façon d'un printf pour . dmrc . Les clés non listées gardent les valeurs précédemm	eri



house (noun)

Haus (noun) deTenTen [2013] freq = <u>7,264,685</u> (364.72 per million)										
Use another candidate translation: erinnern Hausarrest Ordnung Plenum hoch daran Parlament kehren mitteilen										
lick on collocates to access reciprocal bilingual search										
<u>96,897</u>	2.10	VerbY+SubstXDat	1.476.115	3.60	subject_of	<u>58,690</u>	1.90	SubstXNom+VerbY	322,711	0.80
1,667	9.04				overlook	244 (6.20			
<u>1,737</u>	8.94	wohnen	<u>31,400</u>	7.14	stand	<u>734</u> (6.12	beherbergen	<u>1,910</u>	5.78
8,502	8.59	fühlen	<u>26,177</u>	6.04	belong	306	6.09	verfügen	<u>16,744</u>	5.49
3.998	8.10	fahren	<u>46,582</u>	5.99	rebuild	135	5.61	brennen	<u>3,018</u>	5.48
853	7.94	kehren	<u>7,991</u>	5.64	date	266	5.52	abbrennen	<u>364</u>	4.83
	7.90	befinden	<u>30,253</u>	5.35	front			schmiegen	<u>384</u>	4.60
	7.58	schicken	<u>9.019</u>	5.34	burn		4.98	erstrahlen	<u>562</u>	4.56
		sitzen	<u>15,913</u>	5.25				befinden	<u>15,624</u>	4.47
		leben	<u>23,884</u>	5.21				bestechen	<u>933</u>	4.33
		kommen	<u>144,447</u>	5.04				säumen	<u>294</u>	4.17
		holen	<u>10,220</u>	4.99	· ·			gruppieren	<u>255</u>	4.14
		eilen	<u>1,891</u>	4.84				liegen	<u>32,608</u>	3.99
2,400	7.00	rennen	2,939	4.82	3011100	107	4.00	einstürzen	187	3.97
	candidat cates to 96.897 1.667 1.737 8.502 3.998 853 935 2.452 650 1.061 1.284 798	96.897 2.10 1.667 9.04 1.737 8.94 8.502 8.59 3.998 8.10 853 7.94 935 7.90 2.452 7.58 6.507 7.44 9.398 8.10 8.51 7.94 9.52 7.90 2.452 7.58 1.061 7.39 1.284 7.25	Candidate translation: erinnern Hausarra 96,897 2.10 VerbY+SubstXDat (bid dat of) 1,667 9.04 Wohnen (bid dat of) 1,737 8.94 wohnen (bid dat of) 3,998 8.10 fahren (bid dat of) 3,998 8.10 schicken (bid dat of) 935 7.90 befinden (bid dat of) 2,452 7.58 schicken (bid dat of) 1,061 7.39 leben (bid dat of) 1,224 7.25 kommen (bid dat of) 1,224 7.25 schicken (bid dat of) 1,224 7.26 kohnen (bid dat of) 1,224 7.25 kohnen (bid dat of) 2,455 7.00 eilen (bid dat of)	candidate translation: erinnern Hausarrest Ordnuns cates to access reciprocal bilingual search 96.897 2.10 VerbY+SubstXDat (bbl dat of) 1.476.115 1.667 9.04 wohnen 31.400 8.502 8.59 fühlen 26.177 3.998 8.10 fahren 46.582 853 7.94 kehren 7.991 955 7.90 befinden 30.253 2.452 7.58 schicken 9.019 650 7.54 sitzen 15.913 1.061 7.39 leben 23.884 1.246 7.02 kohn 10.220 2.456 7.00 eilen 1.891	Gamma Control Hausarrest Ordnung Plen 96,897 2.10 VerbY+SubstXDat 1.476,115 3.60 1.667 9.04 wohnen 31.400 7.14 8.502 8.59 fühlen 26,177 6.04 3.998 8.10 fahren 46,582 5.95 935 7.94 befinden 30.265 5.35 2.452 7.58 schicken 9.019 5.34 550 7.54 sizzen 1.5.913 5.25 1.61 7.39 leben 2.3.84 5.21 1.284 7.25 kommen 1.4444 5.04 9.935 7.94 leben 2.3.84 5.21 1.616 7.39 leben 2.3.84 5.21 1.284 7.25 kommen 1.44.447 5.04 9.92 7.12 holen 1.891 4.84	Sector Sector<	Sector Sector<	Sector Sector<	addidate translation: erinnern Hausarrest Ordnung Plenum hoch dara Parlament kehren mittellen kates to access reciprocal blingual search 96.897 2.10 Verby+SubstX0at (obi dat of) 1.476.115 3.60 1.667 9.04 (bib dat of) 1.476.115 3.60 overlook 244 6.20 (subj. of) 1.737 8.94 wohnen 31.400 7.14 stand 7.34 6.12 beherbergen 3.998 8.10 fahren 46.582 5.99 1.90 SubstXNom+VerbY 93.5 7.90 befinden 30.253 5.35 behong 306 6.09 verfügen 2.452 7.58 schicken 9.019 5.34 burn 113 4.98 erstrahlen 1.061 7.39 leben 2.3.84 5.21 burn 113 4.98 befinden 7.84 Kommen 1.0.220 4.99 setchen colapse 63 7.1 säumen 9.245 7.20 holen 1.891 4.84 survive 107 4.56 gruppieren	Subject of \$8,690 1.00 Subject of \$8,690 1.90 S

adj_subject_of	<u>7,821</u>	2.10	modifier	<u>174,914</u>	1.30	modifier	<u>6,611,554</u>	1.60	modifies	<u>54,780</u>	0.40
uninhabited	<u>65</u>	7.80	manor	2,323	8.57	unsre	<u>158,009</u>	7.18	price	<u>5,868</u>	8.28
adjoining			guest	<u>2,571</u>	8.06	weiß	<u>40.617</u>	6.84	builder	<u>683</u>	7.85
repossessed	34	7.02	public	4.602	7,74	Ihnen	38,740	6.52	arrest	<u>608</u>	7.82

Create parallel corpora from TMX file

Corpus name	
Source TMX file	Choose File No file chosen See <u>the Wikipedia article on TMX</u>
	Cancel Upload



Terminology extraction

Environment: Extracted keywords and terms

Change extraction options Download keywords: TBX CSV. Download terms: TBX CSV.

Keywords		Score	F	RefF	Terr	ns		Score	F	RefF
C02	W	176.40	<u>12.394</u>	<u>0</u>	c	limate change	W	39.64	<u>54,341</u>	238,935
biodiversity	W	34.38	<u>14.663</u>	65,693	🔲 g	reenhouse gas	W	32.65	11,431	<u>51,682</u>
ecosystems	W	32.71	<u>11.893</u>	54,163	🗌 V	vater quality	W	29.19	<u>9.823</u>	49,251
emissions	W	31.24	<u>54.232</u>	306.028	🗌 c	arbon dioxide	W	26.07	<u>13,115</u>	<u>79,874</u>
📄 unep	W	30.96	<u>3.231</u>	<u>6.603</u>	🗌 n	enewable energy	W	24.73	<u>16.926</u>	<u>113,194</u>
watershed	W	28.70	<u>10,555</u>	<u>54.983</u>	📃 s	ea ice	W	22.66	<u>2.824</u>	<u>10,48</u>
deforestation	W	28.06	<u>5,200</u>	21,498	🔲 g	lobal warming	W	22.15	<u>17,102</u>	129,35
climate	W	27.56	<u>116.973</u>	766.520	🗌 g	lobal climate	W	22.11	<u>3.467</u>	<u>16.40</u>
biomass	W	26.71	<u>9.341</u>	<u>51.698</u>	📃 fo	ossil fuel	W	20.77	<u>4.052</u>	23,47
habitats	W	26.53	<u>9.606</u>	<u>53.974</u>	🗌 S	ustainable development	W	20.64	<u>6.099</u>	<u>41,89</u>
wetlands	W	26.47	<u>9.030</u>	<u>50,123</u>	🗌 c	lean energy	W	19.57	<u>4,694</u>	<u>31,73</u>
greenhouse	W	26.14	22.514	<u>145,622</u>	📃 a	ir pollution	W	17.53	<u>3,941</u>	29,03
desertification	W	25.45	2.448	<u>5.194</u>	🗌 V	vater management	W	16.23	2,222	12,96
🔲 wwf	W	25.25	<u>3.978</u>	<u>16.457</u>	📃 la	and use	W	15.98	<u>4.729</u>	<u>42.16</u>
dioxide	W	24.76	<u>15.611</u>	<u>103.281</u>	📃 k	ow carbon	W	15.75	<u>2.137</u>	<u>12,75</u>
renewable	W	24.24	<u>31.170</u>	223,599	📃 h	iuman health	W	15.69	<u>3.416</u>	27,81
📃 redd	W	23.86	<u>2.938</u>	<u>10,169</u>	0	irganic matter	W	15.68	2.364	<u>15,53</u>
wetland	W	23.63	<u>5.226</u>	28.175	🗌 c	oal-fired power	W	15.23	<u>1.654</u>	7.81
📃 ghg	W	23.52	<u>3.930</u>	<u>18.255</u>	🗌 g	lobal climate change	W	14.90	<u>1.681</u>	8.62
carbon	W	23.37	<u>65,299</u>	<u>500,429</u>	📃 s	olar energy	W	14.71	<u>6.211</u>	65,41
conservation	W	23.14	<u>42.635</u>	<u>325,767</u>	🗌 e	nergy efficiency	W	14.59	<u>6.788</u>	73,28
stormwater	W	22.93	<u>4.148</u>	<u>20,798</u>	🗌 e	environmental impact	W	14.47	<u>3.649</u>	34,23
ecological	W	22.59	<u>14.692</u>	<u>106.963</u>	. v	vind power	W	14.43	3.525	32.76
pollution	W	22.47	<u>27.926</u>	<u>215.718</u>	. v	vater use	W	14.40	<u>2.162</u>	15,49
i forests		22.09	<u>20.339</u>		📄 fr	ood security		14.30	<u>2.726</u>	22.94
sediment		22.04			🔲 s	urface water		14.23		

sketchengine.c

- languages: Czech, Dutch, English, French, German, Chinese (Simplified, Traditional), Italian, Japanese, Korean, Polish, Portuguese, Russian, Spanish
- extraction of terms in source and target languages
- co-occurrence statistics applied
- TBX export
- Wordlist > left sub-menu > language

L1 term	🗢 L2 term	🗘 Logdice 🗢	Co- ≑ freq	L1 freq	L2 freq
prevalence	prévalence	-0.0257005103	306	<u>316</u>	<u>307</u>
soap	savon	-0.0580571016	207	220	<u>211</u>
survival	survie	-0.0683060134	165	<u>170</u>	<u>176</u>
education	éducation	-0.0705785710	1815	<u>1968</u>	<u>1844</u>
adolescence	adolescence	-0.0711610289	89	<u>91</u>	<u>96</u>
condom	préservatif	-0.0840642648	125	<u>139</u>	<u>126</u>
primary prevention	prévention primaire	-0.0840642648	25	27	<u>26</u>
chronological age	âge chronologique	-0.0848888976	33	<u>36</u>	<u>34</u>
basic information	informations de base	-0.0874628413	16	<u>17</u>	<u>17</u>
acid	acide	-0.0874628413	16	<u>17</u>	<u>17</u>
rotavirus	rotavirus	-0.0931094044	15	<u>16</u>	<u>16</u>
universal access	accès universel	-0.0981803939	142	<u>151</u>	<u>153</u>
international guidance	directives internationales	-0.0995356736	14	<u>15</u>	<u>15</u>
stigma	stigmatisation	-0.1040724541	127	<u>133</u>	<u>140</u>
fish	poisson	-0.1043366598	20	<u>21</u>	<u>22</u>
pregnancy	grossesse	-0.1059334447	210	<u>230</u>	222
alcohol	alcool	-0.1110313124	25	<u>28</u>	<u>26</u>
vol	vol	-0.1168136650	83	<u>87</u>	<u>93</u>
syphilis	syphilis	-0.1233824155	28	<u>32</u>	<u>29</u>
public health	santé publique	-0.1235746851	123	<u>133</u>	<u>135</u>
-0	handiaan ())))	0 1007050004	400	400	110

- morphology-aware processing
- our language pipelines for terminology extraction
- TMX, XLIFF, TBX support
- output: terminology inconsistency report
- simple interface
- right now available for testing & evaluation

Translation terminology checking

Upload documents XLIFF or TMX Browse... No file selected. TBX Browse... No file selected. Check POS Check terminology

XLIFF (or TMX) and TBX data must be compatible (= must contain the same language pairs and locales). The files must be in UTF-8 encoding. Use .XLIFF (.xliff) or .TMX (.tmx) suffixes for the filenames.

Your data will not be used in any other way than is required for providing terminology checking feedback on this webpage. Data will be stored only for a few days and will be deleted automatically afterwards.

Terminology Checking: Report

German <=> English, segments: 545, issues: 33, Show all segments

10	Die Sprache der Benutzeroberfläche (nach dem Login) hängt von den Profileinstellungen ab.	The language of the user interface (once you have logged in) depends on the profile settings.	AB => order confirmaton
45	Den Suchbegriff können Sie als <u>Termvorschlag</u> einreichen (Pos.).	You can submit the search term as a term candidate (item).	Termvorschlag $=>$ term suggestion
47	Bitte beachten Sie, dass für das Einreichen eines Termvorschlags keine Wildcards wie * oder % im Suchbegriff enthalten sein dürfen!	When submitting a term candidate, please note that the search term must not contain any wildcards such as * or %.	Termvorschlag => term suggestion
103	Mit einem Klick auf das Diskettensymbol schließen Sie das Hinzufügen einer Definition <u>ab</u> .	Click on the floppy disk symbol when you have finished adding the definition.	AB => order confirmaton
108	In diesem <u>Kontext</u> beziehen sich die Bearbeitungsschritte jedoch auf die Benennung, nicht auf den Begriff!	However, please be aware that here the processing steps refer to the term, not the concept.	Kontext => context
194	Diese Anzeigemöglichkeit hängt jedoch von den für Sie festgelegten Benutzerrechten <u>ab</u> .	However, this display option depends on which user rights you have been assigned.	AB => order confirmaton
282	Die Termdatenbankübersicht listet die aktuell eingerichtete(n) Termdatenbank(en) auf – hier: Interne Terminologie (Pos.).	The Overview of term database lists the term database(s) that is/are currently set up - here: Interne Terminologie (item).	Terminologie => terminology
314	Schließen Sie ihre Eingaben über die Schaltfläche Attribut hinzufügen (Pos.) <u>ab</u> .	When you have finished making your entries, click on Add attribute (item).	AB => order confirmaton
317	Tragen Sie einen Namen ein (Pos.) und schließen Sie Ihre Eingaben über die Schaltfläche Vorlage erstellen (Pos.) ab.	Enter a name (item) and when you have finished making your entries, click on Create template (item).	AB => order confirmaton
511	Stoppwort	Stopword	Stoppwort => forbidden word
512	Wörter, die zum Aussortieren von <u>Termvorschlägen</u> verwendet werden.	Words that are used to filter out term candidates.	Termvorschlag $=>$ term suggestion
525	Ein Wort, das beispielsweise aus einem Termextrakt stammt oder als <u>Termvorschlag</u> eingereicht wurde, ist ein Termkandidat.	For example, a word that originates from a term extract or was submitted as a suggestion is a term candidate.	Termvorschlag => term suggestion
536	Alles was <u>ab</u> "Übersetzen" kommt, passiert in einem Folgeprojekt, in dem dann auch die Ansicht bei der Bearbeitung im Projekt selbst eine zweisprachige Ansicht zur Verfügung steht.	Everything that comes after "Translate" takes place in a follow-up project. The display for processing the tasks in the project is then a bilingual display.	AB => order confirmaton

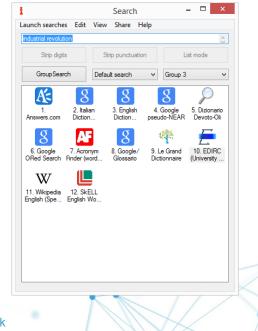
10/16



- from the author's webpage: "IntelliWebSearch is designed to save translators, interpreters, editors and terminologists time when searching the web."
- handy plugin for Windows
- search highlighted text in any application

intelliwebsearch

- from the author's webpage: "IntelliWebSearch is designed to save translators, interpreters, editors and terminologists time when searching the web."
- handy plugin for Windows
- search highlighted text in any application
- configurations for SkE available
- http://ske.li/ske_iws_conf
- planned: configurations for all user corpora



- to our knowledge, the largest parallel corpus from EU resources
- EUR-Lex: http://eur-lex.europa.eu
- "access to European Union law"





- to our knowledge, the largest parallel corpus from EU resources
- EUR-Lex: http://eur-lex.europa.eu
- "access to European Union law"
- 750 million words in English
- 24 languages
- tens of millions of aligned paragraphs
- available for researchers (CC-BY-NC)

- to our knowledge, the largest parallel corpus from EU resources
- EUR-Lex: http://eur-lex.europa.eu
- "access to European Union law"
- 750 million words in English
- 24 languages
- tens of millions of aligned paragraphs
- available for researchers (CC-BY-NC)
- oral presentation at LREC Session: O28 Corpus Querying and Crawling Thursday, 15.55–16.15

Comparison with other corpora (English)

Corpus	Size [M]	Lexicon	Ls	Format
JRC-Acquis	55	N/A	22	XML
DCEP	103	1.0 M	24	TXT
DGT-TM	74	342 k	24	ТМХ
Europarl	61	139 k	21	XML
EUR-Lex	845	2.8 M	24	various

Document type	Author	EuroVoc	Year
Written question	156,744 EC	150,545 EMPTY	142,821 2013 24,978
Regulation	59,758 EP	104,323 EC	15,855 2011 24,852
judicial information	36,964 Provisional data	53,230 EU Member State	13,551 2012 22,879
Decision	20,400 Council of the EU	31,453 import	12,296 2010 22,266
Question at Question Time	19,027 Court of Justice	22,397 Italy	12,240 2007 20,216
Communication	16,384 Court of Justice of the E	U 14,637 Spain	12,206 2008 19,238
Consolidated text	16,060 Court of First Instance	12,201 State aid	12,111 2009 18,088
decision w/out addressee	13,718 General Court	9,056 information transfer	10,946 2006 17,822
Judgment	13,709 EES Committee	4,524 United Kingdom	10,870 2003 16,587
Proposal for a regulation	8,608 United Kingdom	3,995 Greece	10,378 2005 16,407
Opinion	7,774 EEA Joint Committee	2,880 Germany	10,170 2000 16,248
National exec. measures	7,745 Civil Service Tribunal	2,830 EU aid	10,038 2001 16,044
Information	7,314 Malta	2,184 appl. of EU law	9,685 1996 15,293
Notice	7,306 The Member States	1,978 environ. protection	9,559 2004 14,974
Adv. General's Opinion	7,155 Ireland	1,729 France	9,545 1998 14,946
Treaty	5,808 National Courts	1,674 export refund	9,068 1997 14,929
Own-initiative resolution	5,460 Committee of the Regior	ns 1,364 third country	8,524 2014 14,868
Report	4,454 European Court of Audit	ors 1,248 human rights	8,440 2002 14,868
Implementing regulation	4,205 The 12 Member States	1,182 econ. concentration	8,42 1995 14,319
proposal for a decision	4,066 EFTA Surveillance Author	prity 985 merger control	8,276 1999 12,667
Info	4,066 European Central Bank	847 award of contract	8,110 1992 10,768
Directive	3,795 KOSTOPOULOS	807 approx. of laws	7,898 1993 9,693
Order	3,407 Others	686 control of State aid	7,872 1986 9,265
Own-initiative report	3,054 Gov. representatives	639 EU programme	7,617 1990 9,259
Opinion proposing amend.	3,039 The 6 Member States	622 annul. of EC decis.	7,578 1985 9,224

- bilingual terminology from comparable corpora
- definition finding for terminology (Vojta K. at LREC P04)
- term hyper-, hyponyms finding
- multi-word (term) thesaurus
- multilingual terminology extraction
- plugins for Trados, MemoQ
- multilingual diachronic analysis
- one-click terminology...