COST-ENeL-WG3-Automatic Knowledge Acquisition for Lexicography

Survey short title: COSTENeLAutomatic Knowledge Acquisition

Survey long title: COST-ENeL-WG3-Automatic Knowledge Acquisition for Lexicography

Question number: 134

Survey is active

Active from: 13.04.2015

Active until: 13.07.2015

Author: tiberius

Edited: tiberius

Date: 26.03.2015

Date: 13.04.2015

Q1 - Please provide your name(s) and surname(s) - separated by semi-colons if more respondents are providing answers together:

Q2 - Please provide the name of the country where you work:

Q3 -   Please provide the name(s) of the institution(s) where you work (e.g. Institute of the Estonian language, Jozef Stefan institute etc.):

Q4 -   Please provide the description of your position(s) (e.g. lexicographer, software developer, manager etc.):

Q5 - Do you or your institution use a form of automatic knowledge acquisition within a lexicographic project(s)?:
If 'yes' is chosen, you will iterate through a list of questions for different types of automatically acquired knowledge.

○ YES

○ NO

IF (3) Q5 = [1] (YES)
Q6 - Do you or your institution use automatic knowledge acquisition for generating a candidate lemma list?

○ YES

○ NO

Q7 -   Within which project do you or your institution use this type of knowledge acquisition? Please provide the name of the project, a URL and/or bibliographic reference of the project.

```
┌──────────────────────────────────────────────────────────────────────┐
│                                                                        │
│                                                                        │
│                                                                        │
│                                                                        │
│                                                                        │
└──────────────────────────────────────────────────────────────────────┘
```

Q8 - Which corpus / corpora do you use for this type of knowledge acquisition?

```
┌──────────────────────────────────────────────────────────────────────┐
│                                                                        │
│                                                                        │
│                                                                        │
│                                                                        │
│                                                                        │
└──────────────────────────────────────────────────────────────────────┘
```

Q9 -   Did you or your institution publish on your work on this type of automatic knowledge acquisition? If so, please provide bibliographic reference and/or URL. If not, please provide a brief description of the methods/techniques used (max. 150 words)

```
┌──────────────────────────────────────────────────────────────────────┐
│                                                                        │
│                                                                        │
│                                                                        │
│                                                                        │
│                                                                        │
└──────────────────────────────────────────────────────────────────────┘
```

Q10 -   Is the automatically acquired knowledge directly integrated in the published dictionary without human intervention?

○ YES

○ NO

IF (6) Q10 = [2] (NO)

Q11 - Is the automatically acquired knowledge ...

○

○ merely manually checked for
   correctness before publication in the dictionary
   intensively processed by the
   lexicographers and meant to support them in their lexicographic work.

IF (7) Q11 = [2] (

intensively processed by the
lexicographers and meant to support them in their
lexicographic work. If so,
please provide a short description

)

Q12 - Please provide a short description of the work of the lexicographers processing the automatically acquired knowledge.

```
┌──────────────────────────────────────────────────────────────────────┐
│                                                                        │
│                                                                        │
│                                                                        │
│                                                                        │
└──────────────────────────────────────────────────────────────────────┘
```

Q13 -   How do the lexicographers judge the quality of the automatically acquired knowledge? (5-point scale)

○ 1 (very bad)

○ 2

○ 3

○ 4

○ 5 (very good)

Q14 - Do you have any additional comments?

```

```

IF (3) Q5 = [1] (YES)
Q15 - Do you or your institution use automatic knowledge acquisition to extract frequency information, e.g. overall lemma frequency information?

○ YES

○ NO

Q16 -   Within which project do you or your institution use this type of knowledge acquisition? Please provide the name of the project, a URL and/or bibliographic reference of the project.

```

```

Q17 - Which corpus / corpora do you use for this type of knowledge acquisition?

```

```

Q18 -   Did you or your institution publish on your work on this type of automatic knowledge acquisition? If so, please provide bibliographic reference and/or URL. If not, please provide a brief description of the methods/techniques used (max. 150 words)

```

```

Q19 -   Is the automatically acquired knowledge directly integrated in the published dictionary without human intervention?

○ YES

○ NO

IF (10) Q19 = [2] (NO)

Q20 - Is the automatically acquired knowledge ...

○

○ merely manually checked for
correctness before publication in the dictionary
intensively processed by the
lexicographers and meant to support them in their lexicographic work. If so,
please provide a short description

IF (11) Q20 = [2] (

**intensively processed by the
lexicographers and meant to support them in their
lexicographic work. If so,
please provide a short description**

)

Q21 - Please provide a short description of the work of the lexicographers processing the automatically acquired knowledge.

```

```

Q22 -   How do the lexicographers judge the quality of the automatically acquired knowledge? (5-point scale)

○ 1 (very bad)

○ 2

○ 3

○ 4

○ 5 (very good)

Q23 - Do you have any additional comments?

```

```

IF (3) Q5 = [1] (YES)

Q24 - Do you or your institution use automatic knowledge acquisition to extract information on form variation e.g. irregular morphology, orthographic variants?

◯ YES

◯ NO

Q25 - Within which project do you or your institution use this type of knowledge acquisition? Please provide the name of the project, a URL and/or bibliographic reference of the project.

```
┌─────────────────────────────────────────────────────────────────┐
│                                                                   │
│                                                                   │
│                                                                   │
│                                                                   │
└─────────────────────────────────────────────────────────────────┘
```

Q26 - Which corpus / corpora do you use for this type of knowledge acquisition?

```
┌─────────────────────────────────────────────────────────────────┐
│                                                                   │
│                                                                   │
│                                                                   │
│                                                                   │
└─────────────────────────────────────────────────────────────────┘
```

Q27 - Did you or your institution publish on your work on this type of automatic knowledge acquisition? If so, please provide bibliographic reference and/or URL. If not, please provide a brief description of the methods/techniques used (max. 150 words)

```
┌─────────────────────────────────────────────────────────────────┐
│                                                                   │
│                                                                   │
│                                                                   │
│                                                                   │
└─────────────────────────────────────────────────────────────────┘
```

Q28 - Is the automatically acquired knowledge directly integrated in the published dictionary without human intervention?

◯ YES

◯ NO

IF (14) Q28 = [2] (NO)

Q29 - Is the automatically acquired knowledge ...

◯

◯ merely manually checked for
correctness before publication in the dictionary
intensively processed by the
lexicographers and meant to support them in their lexicographic work. If so,
please provide a short description

IF (15) Q29 = [2] (

**intensively processed by the
lexicographers and meant to support them in their
lexicographic work. If so,
please provide a short description**

)

Q30 - Please provide a short description of the work of the lexicographers processing the automatically acquired knowledge.

```

```

Q31 -   How do the lexicographers judge the quality of the automatically acquired knowledge? (5-point scale)

○ 1 (very bad)

○ 2

○ 3

○ 4

○ 5 (very good)

Q32 - Do you have any additional comments?

```

```

IF (3) Q5 = [1] (YES)

Q33 - Do you or your institution use automatic knowledge acquisition to extract example sentences (cf. Vienna workshop)?

○ YES

○ NO

Q34 - **Within which project do you or your institution use this type of knowledge acquisition? Please provide the name of the project, a URL and/or bibliographic reference of the project.**

Q35 - Which corpus / corpora do you use for this type of knowledge acquisition?

Q36 - **Did you or your institution publish on your work on this type of automatic knowledge acquisition? If so, please provide bibliographic reference and/or URL. If not, please provide a brief description of the methods/techniques used (max. 150 words)**

Q37 - **Is the automatically acquired knowledge directly integrated in the published dictionary without human intervention?**

○ YES

○ NO

IF (18) Q37 = [2] (NO)
Q38 - Is the automatically acquired knowledge ...

○

○ merely manually checked for
correctness before publication in the dictionary
intensively processed by the
lexicographers and meant to support them in their lexicographic work. If so,
IF (19) Q38 = [2] (please provide a short description

intensively processed by the
lexicographers and meant to support them in their
lexicographic work. If so,
please provide a short description

)
Q39 - Please provide a short description of the work of the lexicographers processing the automatically acquired knowledge.

Q40 -   How do the lexicographers judge the quality of the automatically acquired knowledge? (5-point scale)

○ 1 (very bad)

○ 2

○ 3

○ 4

○ 5 (very good)

Q41 - Do you have any additional comments?

```


```

IF (3) Q5 = [1] (YES)

Q42 - Do you or your institution use automatic knowledge acquisition to extract multiword expressions (i.e. sequences of words with some unpredictable properties such as "to count somebody in" or "to take a haircut", ranging from collocations and phrasal verbs, (pragmatic) frozen expressions (e.g. of course, good morning) to traditional idioms, proverbs etc.)?

○ YES

○ NO

Q43 -   Within which project do you or your institution use this type of knowledge acquisition? Please provide the name of the project, a URL and/or bibliographic reference of the project. Please also specify what type of multiword expression you extract.

```


```

Q44 - Which corpus / corpora do you use for this type of knowledge acquisition?

```


```

Q45 - Did you or your institution publish on your work on this type of automatic knowledge acquisition? If so, please provide bibliographic reference and/or URL. If not, please provide a brief description of the methods/techniques used (max. 150 words)

```


```

Q46 - Is the automatically acquired knowledge directly integrated in the published dictionary without human intervention?

○ YES

○ NO

IF (22) Q46 = [2] (NO)

Q47 - Is the automatically acquired knowledge ...

○

○ merely manually checked for
correctness before publication in the dictionary
intensively processed by the
lexicographers and meant to support them in their lexicographic work. If so,
please provide a short description

IF (23) Q47 = [2] (

**intensively processed by the
lexicographers and meant to support them in their
lexicographic work. If so,
please provide a short description**

)

Q48 - Please provide a short description of the work of the lexicographers processing the automatically acquired knowledge.

```


```

Q50 - How do the lexicographers judge the quality of the automatically acquired knowledge? (5-point scale)

○ 1 (very bad)

○ 2

○ 3

○ 4

○ 5 (very good)

Q51 - Do you have any additional comments?

IF (3) Q5 = [1] (YES)

Q52 - Do you or your institution use automatic knowledge acquisition to extract neologisms?

○ YES

○ NO

Q53 -   Within which project do you or your institution use this type of knowledge acquisition? Please provide the name of the project, a URL and/or bibliographic reference of the project.

Q54 - Which corpus / corpora do you use for this type of knowledge acquisition?

Q55 -   Did you or your institution publish on your work on this type of automatic knowledge acquisition? If so, please provide bibliographic reference and/or URL. If not, please provide a brief description of the methods/techniques used (max. 150 words)

Q56 -   Is the automatically acquired knowledge directly integrated in the published dictionary without human intervention?

○ YES

○ NO

IF (26) Q56 = [2] (NO)

Q57 - Is the automatically acquired knowledge ...

○

○ merely manually checked for
   correctness before publication in the dictionary
   intensively processed by the

IF (27) Q57 = [2] (lexicographers and meant to support them in their lexicographic work. If so,
   please provide a short description

**intensively processed by the
lexicographers and meant to support them in their
lexicographic work. If so,
please provide a short description**

)

Q58 - Please provide a short description of the work of the lexicographers processing the automatically acquired knowledge.

```

```

Q59 - How do the lexicographers judge the quality of the automatically acquired knowledge? (5-point scale)

○ 1 (very bad)

○ 2

○ 3

○ 4

○ 5 (very good)

Q60 - Do you have any additional comments?

```

```

IF (3) Q5 = [1] (YES)

Q61 - Do you or your institution use automatic knowledge acquisition to extract definitions?

○ YES

○ NO

Q62 -   Within which project do you or your institution use this type of knowledge acquisition? Please provide the name of the project, a URL and/or bibliographic reference of the project.

Q63 - Which corpus / corpora do you use for this type of knowledge acquisition?

Q64 -   Did you or your institution publish on your work on this type of automatic knowledge acquisition? If so, please provide bibliographic reference and/or URL. If not, please provide a brief description of the methods/techniques used (max. 150 words)

Q65 -   Is the automatically acquired knowledge directly integrated in the published dictionary without human intervention?

○ YES

○ NO

IF (30) Q65 = [2] (NO)
Q66 - Is the automatically acquired knowledge ...

○

○ merely manually checked for
   correctness before publication in the dictionary
   intensively processed by the
   lexicographers and meant to support them in their lexicographic work. If so,
IF (31) Q66 = [2] (
   please provide a short description

intensively processed by the
lexicographers and meant to support them in their
lexicographic work. If so,
please provide a short description

)
Q67 - Please provide a short description of the work of the lexicographers processing the automatically acquired knowledge.

Q68 - How do the lexicographers judge the quality of the automatically acquired knowledge? (5-point scale)

○ 1 (very bad)

○ 2

○ 3

○ 4

○ 5 (very good)

Q69 - Do you have any additional comments?

```

```

IF (3) Q5 = [1] (YES)

Q70 - Do you or your institution use automatic knowledge acquisition to extract translation equivalents?

○ YES

○ NO

Q71 - Within which project do you or your institution use this type of knowledge acquisition? Please provide the name of the project, a URL and/or bibliographic reference of the project.

```

```

Q72 - Which corpus / corpora do you use for this type of knowledge acquisition?

```

```

Q73 - Did you or your institution publish on your work on this type of automatic knowledge acquisition? If so, please provide bibliographic reference and/or URL. If not, please provide a brief description of the methods/techniques used (max. 150 words)

```

```

Q74 -   Is the automatically acquired knowledge directly integrated in the published dictionary without human intervention?

○ YES

○ NO

IF (34) Q74 = [2] (NO)

Q75 - Is the automatically acquired knowledge ...

○

○ merely manually checked for
correctness before publication in the dictionary
intensively processed by the
lexicographers and meant to support them in their lexicographic work. If so,
please provide a short description

IF (35) Q75 = [2] (

intensively processed by the
lexicographers and meant to support them in their
lexicographic work. If so,
please provide a short description

)

Q76 - Please provide a short description of the work of the lexicographers processing the automatically acquired knowledge.

```



```

Q77 -   How do the lexicographers judge the quality of the automatically acquired knowledge? (5-point scale)

○ 1 (very bad)

○ 2

○ 3

○ 4

○ 5 (very good)

Q78 - Do you have any additional comments?

```



```

IF (3) Q5 = [1] (YES)

Q79 - Do you or your institution use automatic knowledge acquisition to extract knowledge rich contexts (i.e. a sort of hybrid of a good dictionary example and a definition in the sense that is extracted from a corpus, illustrates the meaning of a term, but it is not a formal definition) ?

○ YES

○ NO

Q80 -   Within which project do you or your institution use this type of knowledge acquisition? Please provide the name of the project, a URL and/or bibliographic reference of the project.

Q81 - Which corpus / corpora do you use for this type of knowledge acquisition?

Q82 -   Did you or your institution publish on your work on this type of automatic knowledge acquisition? If so, please provide bibliographic reference and/or URL. If not, please provide a brief description of the methods/techniques used (max. 150 words)

Q83 -   Is the automatically acquired knowledge directly integrated in the published dictionary without human intervention?

○ YES

○ NO

IF (38) Q83 = [2] (NO)

Q84 - Is the automatically acquired knowledge ...

○

○ merely manually checked for
   correctness before publication in the dictionary
   intensively processed by the
   lexicographers and meant to support them in their lexicographic work. If so,
   please provide a short description

IF (39) Q84 = [2] (

**intensively processed by the
lexicographers and meant to support them in their
lexicographic work. If so,
please provide a short description**

)

Q85 - Please provide a short description of the work of the lexicographers processing the automatically acquired knowledge.

```

```

Q86 -   How do the lexicographers judge the quality of the automatically acquired knowledge? (5-point scale)

○ 1 (very bad)

○ 2

○ 3

○ 4

○ 5 (very good)

Q87 - Do you have any additional comments?

```

```

IF (3) Q5 = [1] (YES)

Q88 - Do you or your institution use automatic knowledge acquisition to extract lexical-semantic relations (e.g. synonyms, antonyms, hypernyms)

○ YES

○ NO

Q89 -   Within which project do you or your institution use this type of knowledge acquisition? Please provide the name of the project, a URL and/or bibliographic reference of the project.

Q90 - Which corpus / corpora do you use for this type of knowledge acquisition?

Q91 -   Did you or your institution publish on your work on this type of automatic knowledge acquisition? If so, please provide bibliographic reference and/or URL. If not, please provide a brief description of the methods/techniques used (max. 150 words)

Q92 -   Is the automatically acquired knowledge directly integrated in the published dictionary without human intervention?

○ YES

○ NO

IF (42) Q92 = [2] (NO)

Q93 - Is the automatically acquired knowledge ...

○

○ merely manually checked for
   correctness before publication in the dictionary
   intensively processed by the
IF (43) Q93 = [2] (
   lexicographers and meant to support them in their lexicographic work. If so,
   please provide a short description

**intensively processed by the
lexicographers and meant to support them in their
lexicographic work. If so,
please provide a short description**

)

Q94 - Please provide a short description of the work of the lexicographers processing the automatically acquired knowledge.

Q95 -   How do the lexicographers judge the quality of the automatically acquired knowledge? (5-point scale)

○ 1 (very bad)

○ 2

○ 3

○ 4

○ 5 (very good)

Q96 - Do you have any additional comments?

┌───┐
│   │
│   │
│   │
│   │
└───┘

 IF (3) Q5 = [1] (YES)
Q97 - Do you or your institution use automatic knowledge acquisition to extract word senses?

○ YES

○ NO

Q98 -   Within which project do you or your institution use this type of knowledge acquisition? Please provide the name of the project, a URL and/or bibliographic reference of the project.

┌───┐
│   │
│   │
│   │
│   │
└───┘

Q99 - Which corpus / corpora do you use for this type of knowledge acquisition?

┌───┐
│   │
│   │
│   │
│   │
└───┘

Q100 -   Did you or your institution publish on your work on this type of automatic knowledge acquisition? If so, please provide bibliographic reference and/or URL. If not, please provide a brief description of the methods/techniques used (max. 150 words)

┌───┐
│   │
│   │
│   │
│   │
└───┘

Q101 - Is the automatically acquired knowledge directly integrated in the published dictionary without human intervention?

◯ YES

◯ NO

IF (46) Q101 = [2] (NO)

Q102 - Is the automatically acquired knowledge ...

◯

◯ merely manually checked for
correctness before publication in the dictionary
intensively processed by the
lexicographers and meant to support them in their lexicographic work. If so,
please provide a short description

IF (47) Q102 = [2] (

**intensively processed by the
lexicographers and meant to support them in their
lexicographic work. If so,
please provide a short description**

)

Q103 - Please provide a short description of the work of the lexicographers processing the automatically acquired knowledge.

```

```

Q104 - How do the lexicographers judge the quality of the automatically acquired knowledge? (5-point scale)

◯ 1 (very bad)

◯ 2

◯ 3

◯ 4

◯ 5 (very good)

Q105 - Do you have any additional comments?

```

```

IF (3) Q5 = [1] (YES)

Q106 - Do you or your institution use automatic knowledge acquisition to extract grammatical patterns (e.g. word profiles, valency)

○ YES

○ NO

Q107 -   Within which project do you or your institution use this type of knowledge acquisition? Please provide the name of the project, a URL and/or bibliographic reference of the project.

Q108 - Which corpus / corpora do you use for this type of knowledge acquisition?

Q109 -   Did you or your institution publish on your work on this type of automatic knowledge acquisition? If so, please provide bibliographic reference and/or URL. If not, please provide a brief description of the methods/techniques used (max. 150 words)

Q110 -   Is the automatically acquired knowledge directly integrated in the published dictionary without human intervention?

○ YES

○ NO

IF (50) Q110 = [2] (NO)

Q111 - Is the automatically acquired knowledge ...

○

○ merely manually checked for
  correctness before publication in the dictionary
  intensively processed by the
  lexicographers and meant to support them in their lexicographic work. If so,
  please provide a short description

IF (51) Q111 = [2] (

**intensively processed by the
lexicographers and meant to support them in their
lexicographic work. If so,
please provide a short description**

)

Q112 - Please provide a short description of the work of the lexicographers processing the automatically acquired knowledge.

```

```

Q113 - How do the lexicographers judge the quality of the automatically acquired knowledge? (5-point scale)

○ 1 (very bad)

○ 2

○ 3

○ 4

○ 5 (very good)

Q114 - Do you have any additional comments?

```

```

Q115 - Do you or your institution use automatic knowledge acquisition to extract linguistic labels (e.g. domain/region/dialect/register/style/time/slang and jargon/ attitude/ offensive terms)?

○ YES

○ NO

Q116 -   Within which project do you or your institution use this type of knowledge acquisition? Please provide the name of the project, a URL and/or bibliographic reference of the project.

Q117 - Which corpus / corpora do you use for this type of knowledge acquisition?

Q118 -   Did you or your institution publish on your work on this type of automatic knowledge acquisition? If so, please provide bibliographic reference and/or URL. If not, please provide a brief description of the methods/techniques used (max. 150 words)

Q119 -   Is the automatically acquired knowledge directly integrated in the published dictionary without human intervention?

○ YES

○ NO

IF (54) Q119 = [2] (NO)

Q120 - Is the automatically acquired knowledge ...

○

○ merely manually checked for
   correctness before publication in the dictionary
   intensively processed by the
   lexicographers and meant to support them in their lexicographic work. If so,
IF (55) Q120 = [2] (
   please provide a short description

**intensively processed by the
lexicographers and meant to support them in their
lexicographic work. If so,
please provide a short description**

)

Q121 - Please provide a short description of the work of the lexicographers processing the automatically acquired knowledge.

Q122 -   How do the lexicographers judge the quality of the automatically acquired knowledge? (5-point scale)

○ 1 (very bad)

○ 2

○ 3

○ 4

○ 5 (very good)

Q123 - Do you have any additional comments?

```
```

IF (3) Q5 = [1] (YES)

Q124 - Do you or your institution use a form of automatic knowledge acquisition not mentioned above.

○ YES

○ NO

Q125 - What type of automatic knowledge acquisition do you use?

```
```

Q126 -   Within which project do you or your institution use this type of knowledge acquisition? Please provide the name of the project, a URL and/or bibliographic reference of the project.

```
```

Q127 - Which corpus / corpora do you use for this type of knowledge acquisition?

```
```

Q128 - Did you or your institution publish on your work on this type of automatic knowledge acquisition? If so, please provide bibliographic reference and/or URL. If not, please provide a brief description of the methods/techniques used (max. 150 words)

Q129 - Is the automatically acquired knowledge directly integrated in the published dictionary without human intervention?

○ YES

○ NO

IF (58) Q129 = [2] (NO)

Q130 - Is the automatically acquired knowledge ...

○

○ merely manually checked for
correctness before publication in the dictionary
intensively processed by the
lexicographers and meant to support them in their lexicographic work. If so,
please provide a short description

IF (59) Q130 = [2] (

intensively processed by the lexicographers and meant to support them in their lexicographic work. If so, please provide a short description

)

Q131 - Click and write question text

Q132 - How do the lexicographers judge the quality of the automatically acquired knowledge? (5-point scale)

○ 1 (very bad)

○ 2

○ 3

○ 4

○ 5 (very good)

Q133 - Do you have any additional comments?

Q134 -   Which types of automatic knowledge acquisition that you or your institution currently do not use would you be most interested in? Please, briefly explain why.

Q135 - Do you have any general comments?