

Description form - PROVIDER

Language Resources: Written Corpus

C.1. GENERAL INFORMATION

Type of resource:

Parallel/Aligned corpus Other:

Creation date:

Update frequency:

Last update:

Application:

Information retrieval Language modeling Machine translation
 Message understanding Natural language processing Topic detection and tracking
 Speech synthesis Other:

C.2. SOURCE DESCRIPTION

Source(s)	Language(s) ¹	Domain(s)	Size (in words, sentences, etc.)	Time span

C.3. ANNOTATION

Linguistic annotation²:

Phonetic Orthographic Morphological Syntactical Semantic
 Other:

Text level of annotation (tagging):

Word Paradigme/Phrase Sentence Other:

Description of the tagging system:

¹ Language codes to be used: **da**=Danish; **de**=German; **en**=English; **es**=Spanish; **fi**=Finnish; **fr**=French; **el**=Greek; **it**=Italian; **nl**=Dutch; **pt**=Portuguese; **sv**=Swedish; ...

² Please, indicate the linguistic annotations in use in section "Further Comments", page 3.

C.4. TECHNICAL INFORMATION

File type:

Text MS Word MS Excel Other:

Workstation:

PC Unix Mac Other:

Format in use:

SGML XML Other:

Standard in use:

ISO/TR 12618:1994 TEI Other:

Character set:

ISO 8859-1 7-bit ASCII 8-bit ASCII Unicode
 Other:

C.5. FURTHER COMMENTS

ELRA
55-57 rue Brillat Savarin
F-75013 PARIS, FRANCE
Tel +33 1 43 13 33 33 - Fax +33 1 43 13 33 30
<http://www.elra.info>