

Corpus of Old Literary Finnish

1 BASIC INFORMATION

1.1 Corpus composition

Written Finnish texts from the years 1543–1809

1.2 Representation of the corpus (flat files, database, markup)

Browsable and searchable on the web

1.3 Character encoding

UTF-8

2 ADMINISTRATIVE INFORMATION

2.1 Contact person

Name: Toni Suutari

E-mail: toni.suutari@kotus.fi

2.2 Copyright statement and information on IPR

Freely browsable and searchable

3 TECHNICAL INFORMATION

3.1 Data structure of an entry

Browsable plain text with a search interface on the web

3.2 Corpus size (num. of tokens)

3 428 618 words

4 CONTENT INFORMATION

4.1 Type of the corpus (monolingual/multilingual, parallel/comparable, raw/annotated)

Monolingual, raw

4.2 The natural language(s) of the corpus

Finnish

4.3 Domain(s)/register(s) of the corpus

Bible translations and religious texts, legal texts, fairy tales, medical books, various.

4.4 Annotations in the corpus (if an annotated corpus)

No annotations

4.5 Intended application of the corpus

Lexicography

4.6 Reliability of the annotations (automatically/manually assigned) – if any

No annotations

5 RELEVANT REFERENCES AND OTHER INFORMATION

http://kaino.kotus.fi/korpus/vks/meta/vks_coll_rdf.xml