

Norsk ordbank

1. BASIC INFORMATION

1.1 *Lexicon type (wordform, explanatory, terminological lexicon, wordnet, etc),*

Wordform with lemma

1.2 *Representation of the lexicon (flat files, database, markup)*

Lines with tab-separated fields

1.3 *Character encoding*

ISO-8859-1, with windows linebreaks (cr/lf)

2. ADMINISTRATIVE INFORMATION

2.1 *Contact person (name, e-mail)*

Arne Martinus Lindstad, ArneMartinus.Lindstad@nb.no

3.2 *Copyright statement and information on IPR*

Available on the condition of signing a form with conditions of use.

Produced by the University of Oslo

3. TECHNICAL INFORMATION

3.1 *Data structure of an entry*

Seven tab-separated fields, described in documentation.

3.2 *Lexicon size (num. of lexical items)*

784240 forms

4. CONTENT INFORMATION

4.1 *The natural language(s) of the lexicon*

Bokmål (Nbo) and Norwegian Nynorsk (Nno)

4.2 *Entry Type*

Wordform

4.3 *Attributes*

Lemma and morphological information, described in documentation

4.4 *Coverage of the lexicon*

The lexicon consists of the following resources:

- a) Word lists and patterns of inflections produced by IBM Norway
- b) Entries and inflection information from *Bokmålsordboka* og *Nynorskordboka* (standard dictionaries for bokmål and nynorsk), produced by ILN, University of Oslo
- c) Codes for argument structure produced by NorKompLeks at NTNU (The Norwegian University of Science and Technology)

4.3 *Intended application of the lexicon*

Language technology

4.4 *Reliability (automatically/manually constructed)*

Manually and automatically constructed

5. RELEVANT REFERENCES AND OTHER INFORMATION

<http://www.nb.no/spraakbanken/tilgjengelege-ressursar/leksikalske-databasar>

<http://www.hf.uio.no/iln/om/organisasjon/edd/forsking/norsk-ordbank/>