

ICAME Web-site

Knut Hofland
The HIT Centre

Up-to-date information on the ICAME archive is found at:

<http://www.hit.uib.no/icame.html>

Here you will find samples and order form for the ICAME CD, corpus manuals in electronic form, back issues of *ICAME Journal*, an archive of the corpora discussion list, the ICAME bibliography and many references to useful resources on the Web.

ICAME – CD-ROM Version 2

<http://www.hit.uib.no/icame/cd>

In addition to the five corpora from the first version, the second version of the ICAME CD-ROM contains an additional 15 corpora. The entire ICAME material is now collected on one single CD-ROM. We no longer distribute single corpora.

The total number of words for all 20 corpora amounts to 17 million (7.9 million words written modern text, 2.5 million transcribed speech and 6.6 million historical text). 3.5 million words have been tagged for part of speech. The modern texts are from the sixties onwards.

The CD comes with a complete version of WordSmith. This program can be used both on the texts on the CD and on the customer's own texts. It can be used for generating and comparing word lists, making concordances and finding collocates and key words for a text.

16 corpora have been indexed with the DOS version of WordCruncher. The search program is included on the CD. WordCruncher can be used directly from the CD. Alternatively the files can be copied to hard disk or network for quicker access.

One of the corpora, COLT, is also indexed with TACT, a DOS program included in the complete version on the CD.

All manuals are included on the CD. They are also available from the ICAME CD home page on the Internet. This homepage also includes an order form, text samples from the individual corpora, program screen display samples and other useful information.

The ICAME CD is available for non-commercial research. Customers will also get access to some of the corpora via the Internet with a web-based interface.

Conditions on the use of ICAME corpus material

The following conditions govern the use of the corpus material distributed through ICAME:

1. No copies of corpora, or parts of corpora, are to be distributed under any circumstances without the written permission of ICAME.
2. Printouts of corpora, or parts thereof, are to be used for bona fide research of a non-profit nature. Holders of copies of corpora may not reproduce any texts, or parts of texts, for any purpose other than scholarly research without getting the written permission of the individual copyright holders, as listed in the manual or on the record sheet accompanying the corpus in question. (For material where there is no known copyright holder, the person(s) who originally prepared the material in computerized form will be regarded as the copyright holder(s).)
3. Commercial publishers and other non-academic organizations wishing to make use of part or all of a corpus or a printout thereof must obtain permission from all the individual copyright holders involved.
4. Publications making use of the material should include a reference to the relevant corpora, giving the name of the corpus and the distributor.

Use of ICAME texts within an institution

Though ICAME texts cannot be used and distributed outside the institution making the order, they can be freely used within the institution (department, faculty, university) for the purposes of research and teaching. To prevent any use of the material for commercial and profit-making purposes, it is advisable to limit access to registered computer users within the institution. The way this is done may vary depending upon the institution making the order.