# DICTIONARIES FORMAT

DICT is a dictionary network protocol created by the DICT Development Group. It is described by RFC 2229. Its goal is to surpass the Webster protocol and to allow clients to access more dictionaries during use. Dict servers and clients use TCP port 2628.

Read the whole articles at <http://en.wikipedia.org/wiki/DICT>.

## FULL DOCUMENTS

* Read the complete specification of this in [RFC2229 - A Dictionary Server Protocol](resources/RFC%202229%20-%20A%20Dictionary%20Server%20Protocol.htm) (HTML).
* A complete details documentation of DICT format can be read at [dictd8](resources/dictd-specification.pdf) (PDF).

## BRIEF & USEFUL INFORMATION

* Databases for dictd are distributed separately. A database consists of two files. One is *a flat text* file, the other *in the index*. The flat text file contains dictionary entries (or any other suitable data), and the index contains tab-delimited tuples consisting of the headword, the byte offset at which this entry begins in the flat text file, and the length of the entry in bytes.
* The offset and length are encoded using base 64 encoding using the 64-character subset of International Alphabet IA5 discussed in RFC 1421 (printable encoding) and RFC 1522 (base64 MIME). Encoding the offsets in base 64 saves considerable space when compared with the usual base 10 encoding, while still permitting tab characters (ASCII 9) to be used for delimiting fields in a record. Each record ends with a newline (ASCII 10), so the index file is human readable.

*ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz0123456789+/*

record C9gA d6

This means: the entry with headword 'record' starts at position C9gA (base 64) in the data file, and it is d6 bytes long. To retrieve the entry, you only need to read d6 bytes from the data file, starting at position C9gA. (Of course, you need to convert these numbers from base 64 to decimal first!)

* Database metadata can be found with entry:
  + *00-database-short*: Shortname of database
  + *00-database-info*: a full description of database
  + *00*-database-url: an url to the database source.