

# Batch Data Upload keywords

---

The first row of the upload Excel/cvs file contains keywords, which can be automatically recognized during the upload as well as names of Properties, names of Conditions and their units.

Two mode are available for upload off data. In the first mode, **scientific format**, data properties to be uploaded are specified as column names. In principle, multiple properties can be allowed to be uploaded in several columns. However, for simplicity we always suggest to upload only one property in one time. The **technical format** is similar to scientific, but it allows to specify property and conditions directly in each row. Thus, it has generic names to indicate that given column is property, condition or are value for them.

The following main **KEYWORDS** are available for **scientific** format (see example [here](#)):

**MOLECULE (SMILE, SMILES, SDF)** is used to upload molecule in one of supported formats

**ARTICLEID (ARTICLE)** indicates articles stored in OCHEM. Names of articles should start with letter "A". The data will be uploaded to the respective article.

**NAME** indicates name of the molecule. Several names can be separated with semi-columns ";" or can be provided in separate columns

**CASRN (CAS)** Chemical Abstract Service number

**N** indicates abbreviation used for molecule in the article, e.g. mol1, M1, cmpA, 13a, 29b **N.B.!** Do not introduce abbreviations as **NAME**.

**UNIT** indicates unit of the property or condition

**PUBMEDID** Articles can be uploaded to OCHEM using RIS format or PubMed id. It can also indicates the respective id in PubMed. The article will be automatically fetched from PubMed if not present in OCHEM.

**COMMENT (COMMENTS)** any additional information about the record

**N.B.!** Properties and conditions are automatically recognized by their names

Additional keywords:

**MOLECULEID** will upload molecule based on its if in OCHEM.

**EXTERNALID** some id in recognized database, e.g. in PubChem, ChEMBL, etc.

**MIXTURE** will allow to upload [MIXTURES](#). See more details in the [MIXTURES](#) section.

**PREDICATE** by default it is equal "=", but also >, <. OCHEM also supports ~, ~=, >=, >>, <<, =< - but they are converted to one of three main ones

**ACCURACY** provides +- intervals of values

**INTERVAL (MAX\_VALUE)** are used to introduce the larger value of the interval values, e.g. 20.1 – 20.3, which are typical for some properties, such as melting point. Notice, that intervals and predicates can be also indicated as part of value to upload.

**PAGE, TABLE, LINE** could be used to indicate the exact place of the value and record in the article

The following three keywords are additionally available for **technical** format (see example [here](#)):

**PROPERTY** specifies that the given cell is the property

**CONDITION** specifies that the given cell is the condition

**VALUE** specifies that the given value is the value for the preceding property or condition

All other previously described KEYWORDS indicated above can be also used to specify meaning of the cells in technical format.