

# [EN] 02. Synonym Dictionary

## The Idea of a Synonym Dictionary

Synonyms are used for increasing the effectiveness of the resource browsing mechanism in digital libraries based on the dLibra system. The dLibra system automatically expands the user's search with words/phrases stored in the synonym dictionary. Therefore, the user can find more publications which may be of interest to him or her. Another reason for creating that mechanism was to make users' searches independent from the conventions of describing digital objects which are used in the given digital library (for example, using various terms for the same concept).

Let us clarify the operating principle of that mechanism with an example. The user writes "Shakespeare" on the author search field. If the synonym dictionary is empty, the query will not be expanded, and only those publications will be found the metadata (or textual content) of which contain the "Shakespeare" string. If, on the other hand, the synonym dictionary defines a synonym, "Szekspir", for "Shakespeare", then the system will be able to expand the search and browse the search index taking into account the Polish spelling of that surname. The user does not have to know which form of the surname was used by the editor when they were describing the publication. It should be emphasized that if the user from the example typed in the "Szekspir" search query, the search would likewise be expanded to include the "Shakespeare" spelling.

Information about concept hierarchies cannot be stored in synonym dictionaries created in the dLibra system.

## The Structure of a Synonym Dictionary

The dLibra system stores a so-called synonym dictionary for every attribute in the metadata schema, in every metadata language. That dictionary contains at least all the values used in the metadata of the resources of the library. The dictionary can also contain words/phrases which are not used directly in the metadata.

Every attribute dictionary for a given language is divided into synonym groups. Every group may contain one or more attribute values. In principle, values belonging to one group should be synonyms, equivalent words, or they should be related in another way.

1. In every group, one group is distinguished and called 'base value'.
2. A group is always named after its base value.
3. Within an attribute dictionary for a given language, there cannot be two or more groups with the same name.
4. Within a group, there cannot be two or more values with the same name. There can be values with the same name for the same attribute in various groups.

Synonym dictionaries can be modified in the Editor Application or in the Administrator Application. In order to add values to a synonym dictionary, the user needs appropriate permissions (not every editor may have them).

Advanced management of synonym dictionaries is possible in the Administrator Application. For example, if a certain number of publications have been ascribed the "Shakespeare" value in the "Author" attribute and no synonym was ascribed to that value, then any number of synonyms can be added easily, without the need to edit every publication separately. If an editor with appropriate permissions adds, at any moment, the "Szekspir" synonym to the "Shakespeare" value, then all publications with the "Shakespeare" value will also be found in response to the "Szekspir" search query.

## Controlled Dictionaries

Every synonym dictionary can be marked as controlled. In such a case, the editor will only be able to use the values entered in that dictionary but not add a new value to it. Controlled dictionaries are useful for attributes which have a limited number of values, for example, the "resource" attribute. It usually does not make sense to use a controlled dictionary for such attributes as "Title" or "Author".