Sort rules

Learn how to configure sort rules for listing records in different domains in the OLIB catalogue. This feature enables you to configure sort rules for different domains, with options to specify various rules such as case sensitivity, character and string replacement, and left-padding numbers with zeros. For example, when sorting a set of bibliographic records by title, numbers spelled out as words (e.g. Nineteen eighty-four) can be “unspelled”, i.e. replaced by the number itself, so that the records sort consistently by number regardless of whether they are entered as numbers or as spelled-out numbers. Sorting rules can be applied to

- Classmarks
- Shelfmarks
- Titles
- Authors

Some predefined rules are also provided, e.g. DDC, UDC have been predefined for Classmarks and Shelfmarks.

Where does the sort configuration have an effect?

The sorting rules are used to generate a sort key for each record in the respective domain. Subsequently, whenever a list of records is sorted by title, author, classmark or shelfmark (e.g. following a title search in the Titles domain, following a search in Folio, or when the contents of a folder are displayed), the records are sorted by this pre-generated sort key rather than being sorted “on the fly”. This includes:

- The sorting of title, author and classmark hitlists in OLIB Web
- The sorting of hitlists in the Copies domain in OLIB Web when the hitlist is sorted by shelfmark
- Other domains’ hitlists that are sorted by title, author, classmark or shelfmark either by default or using the Add More Sorting button in OLIB Web (for example, the Subscription Order Items search in the Order Items domain)
- Search results hitlists in Folio when they are sorted by title, author or classmark [sorting by shelfmark is not an option in Folio]
- Any export Folders (Titles)

In particular Sorting Rules improve the usability of OLIB Stocktaking for libraries which do not use the Dewey Decimal Classification system or similar. (Dewey Decimal Classification system generally sorts correctly without requiring special processing). When sorting rules are enabled, this also affects the sorting of hitlists for the above data, and the use of the configurable Sorting feature for those hitlists.

In order to use the Words To Numbers facility, the language of cataloguing must be assigned to the title record. This can be defaulted at system level or from the cataloguer’s Location for new titles.

The Sorting Rules domain can be found in the Menu under System Administration and is supplied with 6 predefined rules.
sets:

- Titles
- Authors
- DDC classmarks
- UDC classmarks
- DDC shelfmarks
- UDC shelfmarks

System language: this information about configurable sorting relates to English systems. If the non-OLIB elements of your system (server, Oracle and/or Tomcat) are configured to utilise another language by default, then the sorting behaviour may be affected.

**View the default rules**

Go to System Administration> Sorting Rules and perform a wildcard (%) search:

-1 Classmark sorting rules for DDC
-2 Shelfmark sorting rules for DDC based shelfmarks
-3 Classmark sorting rules for UDC
-4 Shelfmark sorting rules for UDC based shelfmarks
-5 Title sorting rules Yes
-6 Author sorting rules

These predefined rules (prefixed by minus sign “-“) cannot be deleted or modified, Apart from Title rules, they are not enabled by default. (NULL is the same as Enabled=No).

When you select a record using the check box, there are 4 Actions supplied with the Sorting Rules domain:

- Apply to applicable records
- Create close copy
- Enable
- Disable

**Apply to applicable records**

This action will enable the rule and commence the update of the appropriate records in the background. For Shelfmark rules this will be applied to all Copies for the Locations associated with the rule or, if no locations are specified, for locations that are not associated with another enabled shelfmark rule. In order to amend this for more specific Copies, there is also an action available in the Copies domain. See below for details.
For Classmarks, the rule will be applied to all classmarks of the select type
• Title and Authors rules will be applied to all titles and authors
• In all cases, if the record is flagged as Manual Sort (see below), the sorting rule set will not be applied to
  that record

Example: apply sorting rules to classmarks (DDC)

1. In System Administration > Sorting Rules, carry out a wildcard search.
2. To apply the default rules, check the box for Classmark sorting rules for DDC.
3. In Other Actions select Apply to Applicable Records.

OLIB responds with a message informing you the sorting rules will be applied to nnn records in the background and to
any new records created.
Click OK and once complete, the Applied Since field will be set to the date and time that the background task has
completed.

Create close copy

This action creates a copy of the selected rule(s), adding an asterisk (*) to the end of the description, to differentiate the
copy from the original in the hitlist. These rules can be modified as required. Use this action to create a new rule set of
your own.
The newly created rule set will not be enabled.

Caution: take care not to enable the rule until all the details are ready.

Example: create a new shelfmark (DDC) rules set

1. In System Administration > Sorting Rules, carry out a wildcard search.
2. To create your new rules record, check the box to select Shelfmark sorting rules for DDC based shelfmarks.
3. In Other Actions select Create Close Copy.

OLIB creates the new record in the hitlist with a positive Sort Rule Number. (Predefined records have a Sort Rule
Number with a minus sign).
Open the record in modify mode and edit as required (field details below).

Disable / enable actions

These actions disable or enable the selected rule(s). This is the only way to enable or disable a predefined rule set, as
predefined rule sets cannot be modified.

Note: the predefined Titles rule set (-5 - Titles sorting rules) cannot be disabled, even using the Disable action.

When you enable a set of rules, it applies those sorting rules to hitlists, and to configurable Sorting, from this point
onwards. If more than one rule is enabled for a record (e.g. the same location is selected for more than shelfmark rule),
then the rule with the highest number will be used.

**Modify a sort rule**

Pre-defined records cannot be modified – modifying applies only to new records.

At the top of the details record is a section which describes and specifies the applicability of this rule:

<table>
<thead>
<tr>
<th>FIELD</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule Type</td>
<td>This determines whether the rule is applied to Authors, Classmarks, Shelfmarks or Titles.</td>
</tr>
<tr>
<td>Description</td>
<td>A description of this rule. This is used by the search for sorting rules in the Sorting Rules domain in OLIB Admin and is shown on the hitlist.</td>
</tr>
<tr>
<td>Class Type</td>
<td>For Classmark rules, this specifies the Class Type to which the rule applies.</td>
</tr>
<tr>
<td>Add Location</td>
<td>For shelfmark rule sets, this is a drop-down list of locations which can be used to add the location codes to the field below, specifying the locations that utilise this sort rule for their shelfmarks.</td>
</tr>
<tr>
<td>Locations</td>
<td>This is a list of the applicable location codes, separated by semicolons (;). You can enter location codes straight into this field rather than selecting them from the drop-down list. If this field is left blank, the shelfmark sorting rule set will be applied to copies at all locations other than those that are specified in an enabled shelfmark rule set. If more than one shelfmark rule set can be applied to a copy, the rule set with the highest sort rule number will be used.</td>
</tr>
<tr>
<td>Enabled</td>
<td>This Yes/No field specifies whether or not the rule is used.</td>
</tr>
<tr>
<td>Applied Since</td>
<td>This is populated with the date and time when the Apply to applicable records action has completed the background processing.</td>
</tr>
</tbody>
</table>

The **Rules for Sorting** section includes the fields that determine the sort rules to be applied. These fields are shown in the order that the feature is applied to the data.
<table>
<thead>
<tr>
<th>FIELD</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific Replacements</td>
<td>The two fields in this section allow for the specific replacement of characters or sequences of characters. These are the first rules to be applied when generating a sort key. They will therefore be processed as case sensitive. Further information on the content of the first field can be found below. The second field is a manually maintained description of the actions in the first.</td>
</tr>
<tr>
<td>Words To Numbers?</td>
<td>If left blank, this will be treated as No. Setting it to Yes will &quot;unspell&quot; a number that is expressed in words, i.e. it will convert the spelled out number to its numeric equivalent. For example, &quot;one&quot; will be converted to &quot;1&quot;. Only “known words” are unspelled. “Known words” are defined in the language record. Refer to the Languages Reference Data section below for a description of how to configure a language’s list of known words and their numeric equivalents. When generating a title sort key, the language record that is referenced is the language of cataloguing that is set in the title record. If the title record does not have a language of cataloguing, spelled out numbers will not be unspelled. The unspelling facility is currently only available for titles. This may subsequently be padded to the left with zeros if the Left-Pad To value stipulates more digits than are generated from the word(s).</td>
</tr>
<tr>
<td>Use Base Letter?</td>
<td>If left blank, this will be treated as No. It is set to Yes in the predefined rule set for titles. This will convert diacritic marks to their base letters. For example, “ë” will become “e”. This is applied after Words To Numbers if Words To Numbers is set to Yes. For example, “una” is Spanish for “1” … “uña” (Spanish for “nail”) will not be converted to “1” because the “ñ” does not become “n” until after the Words To Numbers processing. Note that enabling the base letter option will also enforce case insensitivity and therefore auto-upper case the entire sort key, regardless of whether the Case Sensitive field is set to Yes. This means that you cannot use both base letter conversion and case sensitive sort key generation.</td>
</tr>
</tbody>
</table>
| Case Sensitive?        | If left blank, this will be treated as No and will convert all
<table>
<thead>
<tr>
<th>FIELD</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treat 1000-2999 As Years?</td>
<td>If left blank, this will be treated as No. So you should only set the field to Yes if you want to implement this option. Note, however, that it is set to Yes by default in the predefined Authors rule set as it provides better readability when viewing the date portion of the sort key for an author.</td>
</tr>
<tr>
<td>Convert Roman Numerals?</td>
<td>If left blank, this will be treated as No. Whether to convert Roman numerals to “normal” numbers, e.g. “vii” becomes “7”. This may subsequently by padded to the left with zeros if the Left-Pad To value stipulates more digits than are generated from the numeral.</td>
</tr>
<tr>
<td>Thousands Separator</td>
<td>The character that is normally used as the thousands separator in numbers larger than 999. In English, this is usually a comma (,). This field has no default as a comma may be used for other purposes in classmarks or shelfmarks.</td>
</tr>
<tr>
<td>Decimal Point Character</td>
<td>The character normally used as the decimal separator. In English, this is usually a full stop (.) and this will be used as the default.</td>
</tr>
<tr>
<td>Left-Pad To</td>
<td>The number of digits that a number must include. Numbers with fewer than the specified number of digits will be left-padded with zeros until the requisite number of digits is reached. If left blank, no zeros will be added. This facilitates sorting by number in alphanumeric sort</td>
</tr>
</tbody>
</table>
### Sort Numbers After Letters?

If left blank, this will be treated as No. Whether to sort numbers after letters. This facility may be the preferred option in the following languages: Arabic, Czech, Danish, Dutch, Finnish, French, German, Greek, Hebrew, Hungarian, Italian, Latin, Lithuanian, Norwegian, Russian, Slovak, Spanish, Swedish and Turkish.

### Sort Before Numbers

A free text field for providing characters that should sort before numbers. For example, to sort <, = and > before numbers instead of between numbers and letters simply enter "<=>" into this field and do not enter any of these characters in the next two fields.

### Sort Between Numbers And Letters

A free text field for providing characters that should sort after numbers and before letters. For example, to sort {, | and } before numbers instead of between numbers and letters simply enter "{{}" into this field and do not enter any of these characters in the previous or next field.

### Sort After Letters

A free text field for providing characters that should sort after letters. For example, to sort <, = and > after letters instead of between numbers and letters simply enter "<=>" into this field and do not enter any of these characters in the previous two fields.

---

The Specific Replacements field comprises of a line-by-line list of replacements to be carried out on the data. Each line comprises of 3 or 4 elements as below:

<table>
<thead>
<tr>
<th>ELEMENT</th>
<th>EXAMPLE</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Occurrence, space</td>
<td>1</td>
<td>This is an optional portion of the line providing a single digit and a space</td>
</tr>
<tr>
<td>ELEMENT</td>
<td>EXAMPLE</td>
<td>DESCRIPTION</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>Regular Expression</td>
<td>.</td>
<td>to indicate which of the occurrences of the regular expression should be replaced. In the predefined rules for UDC, the line described in this table (1 .=&gt;) indicates that the first, and only the first, dot should be replaced with a greater than symbol (&gt;). In order to replace a digit and a space with something, you can start the line with * * (or &quot;0 &quot; – a zero and a space) to indicate that all occurrences of that expression should be replaced.</td>
</tr>
<tr>
<td>Equals</td>
<td>=</td>
<td>This specifies an expression to identify what should be replaced. For additional visual clarity and to avoid issues with copy/paste, an underscore (_) will be treated as a space. The hover text for this field lists the characters which must be preceded by a \ to be treated literally. These are: $ ( ) * + &lt; &gt; ? [ ] ^ _ . Shortcuts for &quot;start of word&quot; and &quot;end of word&quot; have been provided as ‘&lt;’ and ‘&gt;’ This is utilised in the author sorting rule to allow to be places “O’Connell” between “Oakley” and “Owen”. Detailed information on regular expressions can be found in the Oracle documentation at: <a href="https://docs.oracle.com/database/121...x.htm#SQLRF020">https://docs.oracle.com/database/121...x.htm#SQLRF020</a></td>
</tr>
<tr>
<td>ELEMENT</td>
<td>EXAMPLE</td>
<td>DESCRIPTION</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>-------------</td>
</tr>
</tbody>
</table>
| Replacement | > | with” elements of the line. When examining the line for this character, the first character (after any occurrence portion) is ignored, allowing the user to specify a replacement for an equals sign by entering “==eq” to replace “=” with “eq”.

Except for replacing an underscore with a space, this portion of the line is read literally as the replacement to use in the resulting sort key. To remove the expression from the sort key completely, simply leave this portion empty. If the regular expression includes any parentheses – ( and ) – to capture the matching portion of the text then the replacement portion of this line can include \1 through to \9 to re-instate the first through ninth captured portions respectively. Using the shortcuts for Start of Word or End of Word will include parentheses. This is utilised in the author sorting rule for handling names such as “O’Connell”.

The field on the right is a manually entered, free text line by line description of the replacement rules stated in the field on the left. This can be useful when trying to interpret replacement lines which look odd at a glance – for example: “$=$”.

Below these fields is a list of ASCII characters in sequence. This is provided for reference to facilitate selection of a suitable character to place before or after a specific other character, e.g:

<table>
<thead>
<tr>
<th>Examples</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>007</td>
<td>007$</td>
</tr>
<tr>
<td>01</td>
<td>01$</td>
</tr>
<tr>
<td>32</td>
<td>32$</td>
</tr>
<tr>
<td>339.5</td>
<td>339&gt;5$</td>
</tr>
<tr>
<td>34</td>
<td>34$</td>
</tr>
<tr>
<td>347</td>
<td>347$</td>
</tr>
<tr>
<td>5</td>
<td>5$</td>
</tr>
<tr>
<td>53+64</td>
<td>53!54$</td>
</tr>
<tr>
<td>53/54</td>
<td>53#54$</td>
</tr>
</tbody>
</table>
Re-sort examples

In this field, simply enter various values line by line, and then click the Re-sort Examples button. This will apply the rules and refresh the page with the lines in the Examples field sorted according to the rules. Once sorted, a second column will be displayed to show the sort key that was generated, so that you can determine why the value is sorted in that position with respect to the other values. For title sort rules, this column also reports the language that was used to unspell a number.

(Note that the columnisation used here is not supported by Internet Explorer.)

<table>
<thead>
<tr>
<th>Examples</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>nineteen eighty-four</td>
<td>0000000019 0000000084</td>
</tr>
<tr>
<td>1984</td>
<td>0000001984</td>
</tr>
<tr>
<td>2001: A Space Odyssey</td>
<td>0000002001: A SPACE ODYSSEY</td>
</tr>
</tbody>
</table>

Note that the columnisation used here is not supported by Internet Explorer. Copying and pasting the value into Excel will present the data in columns should this make the result clearer.

Stocktaking module

Note that the Misfiled report will now take account of the configured sorting rules for the shelfmarks involved in the stocktaking project.

On the Stocktaking project record is a new field to select the sorting rule. This field is applied to the Shelfmark range fields in order to advise the stocktaking process of the range of shelfmarks which are included in the project. The misfiled report will be based on the rules associated with the copy records themselves (if any).

Update shelfmark sorting rules

When at least one Shelfmark rule is enabled a new Other Action is available in the Copies domain: Update shelfmark sorting rules.

When copy records are displayed in the hit list and you check the box to select a copy, this action is available to apply to the copies you have selected. OLIB displays the available options you can choose from in an alert box.

https://help.oclc.org/Library_Management/OLIB/Catalogue/Maintain_the_catalogue/Sort_rules
Printed: Wed, 02 Mar 2022 19:43:02 GMT
<table>
<thead>
<tr>
<th>OPTION</th>
<th>PURPOSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shelfmark sorting rules for DDC based shelfmarks</td>
<td>This will assign the selected sorting rule to the selected copies, by adding the selected rule set’s ID to the copy’s Sort Rules Used field. This assignment will be overridden if another Shelfmark sorting rule is applied from the Sorting Rules domain to all copies at their location(s).</td>
</tr>
<tr>
<td>Remove Configured Sorting</td>
<td>This will remove the assignment of a sorting rule from the selected copy records. The sort key will be removed, returning the sorting of the selected copies to the default – UPPER(shelfmark).</td>
</tr>
<tr>
<td>Set Sorting To Automatic</td>
<td>This will set the Manual Sort Key? setting to NULL, assuring that the generated sort value is effective.</td>
</tr>
<tr>
<td>Set Sorting To Manual</td>
<td>This will set the Manual Sort Key? setting to Yes, preventing any further automatic updates of the generated sort value.</td>
</tr>
</tbody>
</table>

Ye thorny issue with ethels

In older publications, there may be English titles which use characters such as an old style long S, for example “Paradije lojč”. With the Sorting Rules configuration these can be placed as equivalent to an “s” in the hit list.

Title sorting and languages reference data

If Words To Numbers is set to Yes in the active Titles sorting rule set, the language of cataloguing specified in the title record will be used to determine which words will be changed to numbers (before being padded with zeros). If the title record does not have a language of cataloguing, the default cataloguing language specified in OLIB Defaults will be used as the language of cataloguing. If this default is not set, word-to-number conversion will not take place.

In support of this, the Words For Numbers sheet has been added to the Cataloguing Reference Data >Languages layout:

More about words for numbers

This sheet is used to maintain the list of translations from a number as a word to a numeric value. English numbers 1 – 100 (cardinal and ordinal) have been added as predefined to the standard “ENG” language record. The Input field is used to list, line by line, a “word=number” conversion, as shown on the saved list to the right. Once entered, the Save Input button can be used to add the items in the Input field to the saved list. The Input field may contain spaces in front of the equals sign (e.g. “one hundredth=100”). The matching will only be applied for entire words (or phrases). For example, “twofold” will not be converted to “2fold”. Items that have just been added to the saved list can be removed immediately by clicking the Remove Input button. Items can be removed from the saved list by selecting them and using the Delete option. Alternatively, use the Remove For Edit action to delete the selected items from the saved list and place them into the Input field. They can then be modified before being re-added to the

https://help.oclc.org/Library_Management/OLIB/Catalogue/Maintain_the_catalogue/Sort_rules
Printed: Wed, 02 Mar 2022 19:43:02 GMT
saved lists using the Save Input button. This provides a quick way to modify an item without having to delete it and re-enter it from scratch, or to copy an item easily. Special cases may need to be considered. For example, without adding a special Words for Numbers rule in the English language record, “Nineteen eighty-four” will be converted to “19 84” and will therefore be sorted between “19” and “20”. Users may expect this to appear between “1983” and “two thousand”. Thus, you may want to consider adding the following in Words For Numbers: nineteen eighty-four=1984

Sort Key Generation and Titles with Subtitles

The title sort key is generated as a composite of the title, the subtitle and the surname of the primary author. If you look at the sort key that has been generated for a record with a title and a subtitle, you will see that the separator between the title and the subtitle is an exclamation mark rather than a colon. This is because the colon comes after the numbers in the ASCII table. Thus, if the colon were to be used as the separator in the sort key, records with subtitles beginning with a number would sort incorrectly in relation to records that have an identical main title except that one of them has a number at the end. For example:

RDA Cataloguing : 101 ways to skin a catalogue

RDA Cataloguing 101 : RDA cataloguing for dummies

If the colon were to be used as the title/subtitle separator in the sort key, the 1st record would sort after the 2nd record in the above example. The exclamation mark is the lowest character in the ASCII table, so using it as the title/subtitle separator in the sort key results in the above records being sorted correctly.

Manual sort

If a sorting rule set is enabled for a domain (Titles, Names, Classmarks or Shelfmarks (in Copies)), a sort key is automatically generated for each record in that domain. It is possible to modify this sort key manually if the configurable rules do not get it quite right. If a sort key is modified manually, you must also set the respective Manual Sort Key? field to Yes, otherwise it will revert to the automatically generated sort key when the record is next saved.

Each domain includes a pair of attributes to enable you to set the sort key manually for a record in that domain:

- **Sort Key** This field contains the sort key that has been generated for the current record. It can be modified, but if it is, you must also set Manual Sort Key? to Yes, otherwise it will revert to the automatically generated sort key when the record is next saved.
- **Manual Sort Key?** This Yes/No field will, if set to Yes, stop any further automatic changes to the sort key.
Flag a title as a “super title”

Using the manual sorting facility, it is now possible to flag a title record as a “super title”, one that is sorted above other records in a hitlist. This can be done simply by adding an exclamation mark (!) at the beginning of the record’s sort key and setting the Manual Sort Key flag to Yes.

Spanish language

If the underlying Oracle system is configured as a Spanish Oracle system, you will need to add an exclamation mark and a space, rather than just an exclamation mark, as the exclamation mark is a reserved character in a Spanish Oracle system.

Leading articles

Leading articles are defined as words included in the Stop Word List that have Non-File set to Yes. Leading articles in titles and authors will be removed from the start of the sort key. This is done automatically and cannot be switched off in the Sorting Rules configuration.

Reports module

To take advantage of the new sorting methods in Reports, the SQL must be changed to use the new sort keys.

In titles this will be “ORDER BY sort_title_main_key”. This value should always be populated and there is no longer a need to consider non_file_chars when using this value as it is already accounted for.

For copies sorted by Shelfmark, this will be “ORDER BY NVL(sort_shelfmark_key, shelfmark)”.

For classes this will be “ORDER BY NVL(sort_key, classmark)”.

For authors this will be “ORDER BY f_sname”, once the appropriate sorting rule has been applied.