

Overview

This is a partial overview of the OpenOffice/LibreOffice panel and the code behind.

- To access the panel: `JabRef:[menu:View]/[OpenOffice/LibreOffice]`
- The user documentation is at <https://docs.jabref.org/cite/openofficeintegration>

I am going to refer to OpenOffice Writer and LibreOffice Writer as LibreOffice or LO: their UNO APIs are still mostly identical, but I only tested with LibreOffice and differences do exist.

Subject

- What is stored in a document, how.
- Generating citation markers and bibliography
 - (excluding the bibliography entries, which is delegated to the layout module)

The purpose of the panel

- Allow the user to insert **citations** in a LibreOffice writer document.
- Automatically format these according to some prescribed style as **citation markers**.
- Generate a **bibliography**, also formatted according to the style.
 - The bibliography consists of a title (e.g. “References”) and a sorted list of formatted bibliography entries, possibly prefixed with a marker (e.g. “[1]”)
- It also allows some related activities: connect to a document, select a style, group (“Merge”) the citations for nicer output, ungroup (“Separate”) them to move or delete them individually, edit (“Manage”) their page-info parts, and collect the database entries of cited sources to a new database.

Citation types

Citations (actually citation groups, see below) have three types depending on how the citation marker is intended to appear in the text:

- **Parenthesized**: “(Smith, 2000)”
- **In-text**: “Smith (2000)”
- **Invisible**: no visible citation mark.
 - An invisible citation mark lets the user to use any form for the citation by taking control (and responsibility) back from the style.

- Like the other two citation types, they have a location in the document.
- In the bibliography these behave as the other two citation types.
- In LibreOffice (`LibreOffice:[Ctrl-F8]` or `LibreOffice:[menu:View]/[Field Shadings]`) shows reference marks with gray background. Invisible citation marks appear as a thin gray rectangle.
- These citation types correspond to `\citep{Smith2000}`, `\citet{Smith2000}` in [natbib](#) and `\nocite{Smith2000}`. I will use `\citep`, `\citet` and `\citen` in “LaTeX pseudocode” below.

PageInfo

The citations can be augmented with a string detailing which part of a document is cited, for example “page 11” or “chapter 2”.

Sample citation markers (with LaTeX pseudocode):

- `\citep[page 11]{Smith2000}` “(Smith, 2000; page 11)”
- `\citet[page 11]{Smith2000}` “Smith (2000; page 11)”
- `\citen[page 11]{Smith2000}` “”
- This string is referred to as `pageInfo` in the code.
- In the GUI the labels “Cite special”, “Extra information (e.g. page number)” are used.

Citation groups

Citations can be grouped.

A group of parenthesized citations share the parentheses around, like this:

“(Smith, 2000; Jones 2001)”.

- Examples with pseudocode:
 - `\citep{Smith2000,Jones2001}` “(Smith, 2000; Jones 2001)”
 - `\citet{Smith2000,Jones2001}` “Smith (2000); Jones (2001)”
 - `\citen{Smith2000,Jones2001}` “”

From the user’s point of view, citation groups can be created by

- 1 Selecting multiple entries in a bibliography database, then

- `[click:Cite]` or
- `[click:Cite in-text]` or
- `[click:Cite special]` or
- `[click:Insert empty citation]` in the panel.

This method allows any of the citation types to be used.

2 `[click:Merge citations]` finds all sets of consecutive citations in the text and replaces each with a group.

- `(change)` The new code only merges consecutive [parenthesized](#) citations.
 - This is inconsistent with the solution used in `[click:Cite]`
 - My impression is that
 - groups of in-text or invisible citations are probably not useful
 - mixed groups are even less. However, with a numbered style there is no visual difference between parenthesized and in-text citations, the user may be left wondering why did merge not work.
 - One way out could be to merge as a “parenthesized” group. But then users switching between styles get a surprise, we have unexpectedly overridden their choice.
 - I would prefer a visible log-like warning that does not require a click to close and lets me see multiple warnings. Could the main window have such an area at the bottom?
- Starting with JabRef 5.3 there is also `[click:Separate citations]` that breaks all groups to single citations.
 - This allows
 - deleting individual citations
 - moving individual citations around (between citation groups)
 - (copy does not work)
 - (Moving a citation within a group has no effect on the final output due to sorting of citations within groups. See [Sorting within a citation group](#))

In order to manage single citations and groups uniformly, we consider each citation in the document to belong to a citation group, even if it means a group containing a single citation.

Citation groups correspond to citation markers in the document. The latter is empty for invisible citation groups. When creating the citation markers, the citations in the group are processed together.

Citation styles

The details of how to format the bibliography and the citation markers are described in a text file.

- These normally use `.jstyle` extension, and I will refer to them as jstyle files.
- See the [User documentation](#) for details.
- I will refer to keywords in jstyle files as `jstyle:keyword` below.

Four major types citation of styles can be described by a jstyle.

- (1) `jstyle:BibTeXKeyCitations`
 - The citation markers show the citationKey.
 - It is not fully implemented
 - does not produce markers before the bibliography entries
 - does not show pageInfo
 - It is not advertised in the [User documentation](#).
 - Its intended purpose may be
 - (likely) a proper style, with “[Smith2000]” style citation markers
 - (possibly) a style for “draft mode” that
 - can avoid lookup of citation markers in the database when only the citation markers are updated
 - can produce unique citation markers trivially (only needs local information)
 - makes the citation keys visible to the user
 - can work without knowing the order of appearance of citation groups
 - In case we expect to handle larger documents, a “draft mode” minimizing work during `[click:Cite]` may be useful.
- There are two types of numbered (`jstyle:IsNumberEntries`) citation styles:
 - (2) Citations numbered in order of first appearance (`jstyle:IsSortByPosition`)
 - (3) Citations numbered according to their order in the sorted bibliography
- (4) Author-year styles

Sorting

Sorting the bibliography

The bibliography is sorted in (author, year, title) order

- except for `jstyle:IsSortByPosition`, that uses the order of first appearance of the cited sources.

Ordering the citations

The order of appearance of citations (as considered during numbering and adding letters after the year to ensure that citation markers uniquely identify sources in the bibliography) is decided on two levels.

- 1 Their order within each citation group (`localOrder`), and
- 2 the order of the citation groups (citation markers) in the text (`globalOrder`).

SORTING WITHIN A CITATION GROUP (`localOrder`)

The order of citations within a citation group is controlled by `jstyle:MultiCiteChronological`.

- `true` asks for (year, author, title) ordering,

- `false` for (author, year, title).
- (There is no option for “in the order provided by the user”).

For author-year citation styles this ordering is used directly.

- The (author, year, title) order promotes discovering citations sharing authors and year and emitting them in a shorter form. For example as “(Smith 2000a,b)”.

For numbered styles, the citations within a group are sorted again during generation of the citation marker, now by the numbers themselves. The result of this sorting is not saved, only affects the citation marker.

- Series of consecutive number are replaced with ranges: for example “[1-5; 11]”

ORDER OF THE CITATION GROUPS (`globalOrder`)

The location of each citation group in the document is provided by the user. In a text with no insets, footnotes, figures etc. this directly provides the order. In the presence of these, it becomes more complicated, see [Order of appearance of citation groups](#).

ORDER OF THE CITATIONS

- `globalOrder` and `localOrder` together provide the order of appearance of citations
- This also provides the order of first appearance of the cited sources.

First appearance order of sources is used

- in `jstyle:IsSortByPosition` numbered styles
- in author-year styles: first appearance of “Smith200a” should precede that of “Smith200b”.

To achieve this, the sources get the letters according the order of their first appearance.

- This seems to contradict the statement “The bibliography is sorted in (author, year, title) order” above.

It does not. As of JabRef 5.3 both are true.

Consequence: in the references Smith2000b may precede Smith2000a. ([reported](#))

- Some author-year citation styles prescribe a higher threshold on the number of authors for switching to “FirstAuthor et al.” form (`jstyle:MaxAuthors`) at the first citation of a source (`jstyle:MaxAuthorsFirst`)

What is stored in a document (JabRef5.2)

- Each group of citations has a reference mark.

(Reference marks are shown in LibreOffice in Navigator, under “References”.

To show the Navigator: `LibreOffice:[menu:View]/[Navigator]` or `LibreOffice:[key:F5]`)

Its purposes:

- 1 The text range of the reference mark tells where to write or update the citation mark.
- 2 The name of the reference mark
 - Lets us select only those reference marks that belong to us
 - Encodes the citation type
 - Contains the list of citation keys that belong to this group
 - It may contain an extra number, to make the name unique in the document
 - Format: `"JR_cite{number}_{type}_{citationKeys}"`, where
 - `{number}` is either empty or an unsigned integer (it can be zero) to make the name unique
 - `{type}` is 1, 2, or 3 for parenthesized, in-text and invisible
 - `{citationKeys}` contains the comma-separated list of citation keys
 - Examples:
 - `JR_cite1_Smith2000` (empty number part, parenthesized, single citation)
 - `JR_cite0_2_Smith2000,Jones2001` (number part is 0, in-text, two citations)
 - `JR_cite1_3_Smith2000,Jones2001` (number part is 1, invisible, two citations)
- Each group of citations may have an associated pageInfo.
 - In LibreOffice, these can be found at `LibreOffice:/[menu:File]/[Properties]/[Custom Properties]`
 - The property names are identical to the name of the reference mark corresponding to the citation group.
 - JabRef 5.2 never cleans up these, they are left around.
(problem) New citations may “pick up” these unexpectedly.
- The bibliography, if not found, is created at the end of the document.
 - The location and extent of the bibliography is the content of the Section named `"JR_bib"`. (In LibreOffice Sections are listed in the Navigator panel, under “Sections”)
 - JabRef 5.2 also creates a bookmark named `"JR_bib_end"`, but does not use it. During bibliography update it attempts to create it again without removing the old bookmark. The result is a new bookmark, with a number appended to its name (by LibreOffice, to ensure unique names of bookmarks).
 - [Correction in new code](#): remove the old before creating the new.

How does it interact with the document?

- “stateless”

JabRef is only loosely coupled to the document.

Between two GUI actions it does not receive any information from LibreOffice.

It cannot distinguish between the user changing a single character in the document or rewriting everything.

- Access data

- During a `[click:cite]` or `[click:Update]` we need the reference mark names.
 - Get all reference mark names
 - Filter (only ours)
 - Parse: gives citation type (for the group), citation keys
 - Access/store pageInfo: based on reference mark name and property name being equal
 - Creating a citation group: (`[click:cite]`)
 - Creates a reference mark at the cursor, with a name as described above.
 - Update (refreshing citation markers and bibliography):
 - citation markers: the content of the reference mark
 - bibliography: the content of the Section (in LibreOffice sense) named `"JR_bib"`.
-