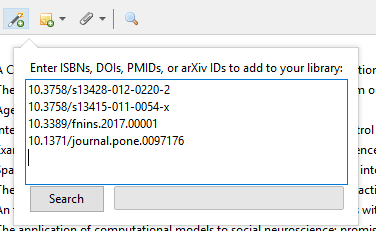
Data extraction guidelines

# PDF files

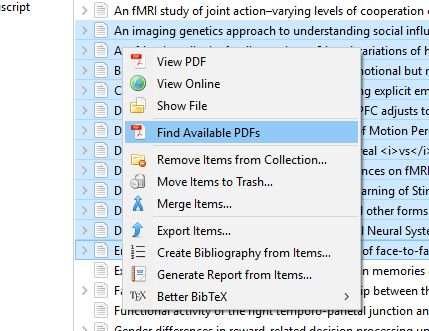
PDF files for published articles can be automatically extracted using the free reference manager ZOTERO, provided that one has access to at least the DOIs of the articles of interest.

## Step by step procedure:

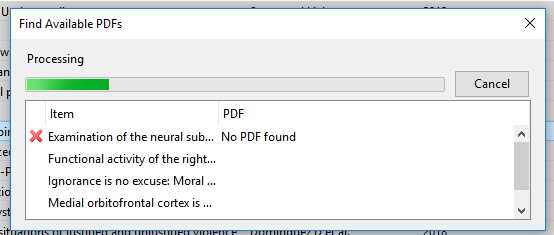
1. Open Zotero
2. Import the meta-data of the articles for which PDFs will be extracted. If data were stored in CSV or other list format, copy the column containing the DOIs of the articles from the list formatted file, and paste into the “add item(s) by identifier” tab in Zotero. Then press “Search”:



1. In Zotero, mark the references to extract PDFs for. To mark all references in the collection, press **ctrl+a**.
2. Right-click on any marked reference. Then click “Find Available PDFs”:

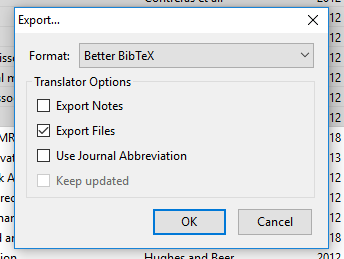


1. A download tracker will appear. Zotero will automatically search for and download available PDFs for all marked references that do not already have a PDF. If there are many PDFs to search for, this process can take a long time. When the search-and-download process has finished, click “close” in the top right corner of the window.



1. PDFs for references that Zotero could not find a PDF for will have to be manually added to the Zotero library. Usually, Zotero should be able to find the PDFs for at least 90% of references. If many fewer than this are found, it may be that publisher paywalls are blocking Zotero from downloading articles. It may help to try again while connected to University Wifi.

Once PDFs have been added for all references, mark all references, right-click on any marked reference, and click “export item”. In the prompt that appears, check the “Export Files” box, then click “OK”. In the next prompt, select where you would like to save the library and click “Save”.



1. Your PDF files exist inside the directory of your saved library. To locate them, go to <your library>/files/