

Google Citation User ID Search Algorithm

Description

The Google Citation User ID Search algorithm is implemented in Java. The Google Citation user ID is a web reader algorithm that retrieves the Google citation for the specified author by querying directly through [Google Author Search](#).

Warning: Google prevents crawling activity on Google Scholar search. To avoid your IP being blocked by Google, it is recommended to search for no more than 100 records per day with this algorithm. However, Google does allow unlimited requests for the remaining Google citation algorithms ([Attach Citation Indices from Google Scholar](#), and [Attach Citation Table from Google Scholar](#)) once the user IDs are known. Limitations on crawling activity for Google Scholar are provided in the <http://scholar.google.com/robots.txt> file.

Given a table with a column of author names as input, the algorithm returns a table with the Google Citation user ID and information associated with each author. Multiple results found will be indicated with same unique index and '*' will be used to indicate the default author. Once the user IDs are retrieved, the following algorithms can be invoked:

- [Attach Citation Indices from Google Scholar](#)
- [Attach Citation Table from Google Scholar](#)

The algorithm takes 2 parameters.

- **Author column** is the name of the column that contains the author's full name. It is recommended to provide the author's name in [First name] [Middle name] [Last name] format with a space as separator.
- **Delimiter** delimits the tokens in the Author column that contain multiple authors per entry. When constructing your tables, do not use a separator that is used as a whole or part of any token.

The results will be generated into a CSV file with the following fields:

- **Citation User ID** is the Google Citation user ID returned by Google Scholar Search.
- **Queried Name** is the author name that was used to query Google Scholar Search.
- **Author** is the author name returned by Google Scholar Search.
- **University** is the university registered under the Google Citation Author profile.
- **Verified email** is the email registered under the Google Citation Author profile.
- **Unique Index** is an identifier number assigned by Sci2 to uniquely identify results from same query.
- **Combine Values** indicates the recommended results for the queries with '*' sign. This is useful if a query returns multiple results.

Please see 'Usage Hints' for more guidance and details.

Pros & Cons

Because of the usage limitation from Google, it is the user's own responsibility to prevent their IP from being black-listed by Google. Once the Google Citation user IDs are retrieved, it is safe to invoke the remaining Google Citation algorithms without limitation.

Applications

The Google Citation User ID algorithm provides a service to retrieve the Google Citation user ID for the given author names. The retrieved Google Citation user ID can be used for querying the specified author h-index, citation information, etc.

Implementation Details

The Google Citation User ID Search algorithm is implemented in Java. Each given author name will be wrapped into a Google Scholar Search query. The returned results will be parsed into a table format. Multiple results could be returned and users are responsible for deleting duplicated results.

Usage Hints

Please read the Description section before continuing to avoid your machine's IP from being black-listed by Google.

Steps:

1. Load the CSV file by choosing Load from menu bar.
2. Choose File > Google Scholar > Google Citation User ID Search
3. A window will pop up and with a list of two input parameters.
4. Select the author column to be processed and the delimiter that separates the author field. (Ignore this if each field contains only one author.)
5. Press OK and a new result will be generated in the csv file which can be seen in the Data Manager (right panel). Save it and you can view the results with Excel.
6. Please refer to the Description for information about the result fields.
7. Now you can use the user IDs to request the author information by invoking [Attach Citation Indices from Google Scholar](#) and [Attach Citation Table from Google Scholar](#).

Links

- [Source Code](#)

See Also



The license could not be verified: License Certificate has expired! [Generate a Free license now.](#)