A Sort-of Simple Method of Using ISI Web of Science to Get All Works-Cited from All Articles By One or More Authors

Jeremy Cusker  
Cornell University

Follow this and additional works at: https://surface.syr.edu/nyscilib

Part of the Library and Information Science Commons

Recommended Citation
Cusker, Jeremy, "A Sort-of Simple Method of Using ISI Web of Science to Get All Works-Cited from All Articles By One or More Authors" (2012). Upstate New York Science Librarians Conference. 19.
https://surface.syr.edu/nyscilib/19

This Presentation is brought to you for free and open access by SURFACE. It has been accepted for inclusion in Upstate New York Science Librarians Conference by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.
A Sort-of Simple Method of Using ISI Web of Science to Get All Works-Cited from All Articles By One or More Authors

Jeremy Cusker
Fall 2012
Have you ever wanted to find out:

*How can I find out what the faculty are really reading?*

Specifically, what are they citing in their own papers?

Even more specifically, is there a way to retrieve a list of all works-cited in ISI Web of Science *by all articles by one (or more) authors?*
Web of Science can provide you with the
-works citing
-works cited by
... a *single* article easily enough.
It’s also easy to search by author . . .

Create a citation report . . .

And use the “Citing Articles” function to see all works that cite all articles by that author.
BUT the opposite case is not true: There is no way to use ISI to search for an author (much less multiple authors) and then find all of the works their papers have cited.

Or so it seemed . . .
It always seemed to me that if we could find all the journals that faculty in a department were citing, that would give us a picture of what they were really reading.

It could also be a really useful tool for collection development.
Step 1:

This is what Cornell’s Earth & Atmospheric Sciences (EAS) department looks like in the form of a Web of Science Author search:

allmendinger r* or allmon w* or andronicos c* or barazangi m* or brown l* or cathles l* or chen g* or cisne j* or colucci s* or degaetano a* or derry l* or green c* or hysell d* or jordan t* or kay r* or kay s* or lohman r* or mahowald n* or phipps-morgan j* or pritchard m* or riha s* or white w* or wilks d* or wysocki m*
Step 2:

Narrow it down by date and then perform the search.

Step 3:

Post-coordinate by using the Institution facet to limit to Cornell.
Step 4:
- Select all records
- Choose Full Record, plus Cited References
- Export to Marked List

279 records matched your query of the 5,267,559 in the data limits you selected.
Step 5: In the Marked List view, select *only* Cited References as a field to export.

Then select Tab-Delimited Format (Win) and Save
Step 6: When prompted, SAVE the record set, *don’t* open it directly from ISI.

If the "Save As" dialog does not appear automatically, then click "Save" to download the records.

When you are done saving the file, click "Return."
Step 7: Now we switch to Excel to open the file.

Choose “Delimited” format in the first prompt.

Choose “Semicolon” (and unchoose “Tab”) in the second prompt.

Click through the rest of the prompts as-is.
Whew, what a mess, huh?

Stay with me: We’re getting close.
After Row A which is record prefixes, the rest of the rows are exactly what you want: Citations of other journal articles.

The problem is, there’s some extra stuff in there you don’t want.

So, start trimming the excess.

Step 8: Open a new tab of the spreadsheet and put this function into the first cell of the first column:

=TRIM(RIGHT(SHEET1!A1,LEN(SHEET1!A1)-FIND("," ,SHEET1!A1,FIND("," ,SHEET1!A1)+1)))
This clears away everything to the left of the second comma in each cell.

...and then drag the formula of that cell down the column to copy that function into all cells.
Step 9: Then Sort, A-to-Z descending.

And there you will have it . . .

Journal titles (abbreviations, anyway), by frequency, in alphabetical order.
Caveats, cheats, and improvements:

1.) You will need to do this (steps 8-9) for every remaining column in the spreadsheet, “pouring” them end-to-end into a single column. That can take a while, but it’s manageable.

2.) There is a certain amount of by-hand cleanup you’ll need to do—a few citations are formatted idiosyncratically.
Caveats, cheats, and improvements, continued:

3.) The more author names and/or the more years of coverage you want to export in step 5, the slower it goes: The record set grows exponentially and is a very crunchy task for WoS. It’s not unusual for it to take 4-5 minutes to export.

4.) Just before Step 8, you may want to first Pivot the table, so that all citations from a single article are in the same column rather than the same row.

You’ll eventually get the same results either way, but the “interim” results will be more immediately meaningful (i.e. you can see right away if individual articles are citing the same journal multiple times).
So, back to our real-world example:

If I run this process with the aforementioned list of authors from the Department of Earth and Atmospheric Sciences, I get these top 15 journals cited by the faculty therein:

- Physical Review. D. Particles, Fields, Gravitation and Cosmology
- Physical Review Letters
- Physics Letters B
- Geophysical Research Letters
- Geochemistry, geophysics, geosystems
- Journal of Geophysical Research
- Annales Geophysicae [Germany]
- Marine and Petroleum Geology
- Geology**
- Journal of Geophysical Research. Space Physics
- Nature
- Science
- Earth and Planetary Science Letters
- Field Crop Research
The significance of this can be seen if we compare these findings with the top journals in geoscience fields as given by Eigenfactor and ISI Categories…

<table>
<thead>
<tr>
<th>ISI Category</th>
<th>ISI Category</th>
<th>Eigenfactor &quot;Science Browser&quot;</th>
<th>Most-cited in faculty papers (via ISI Web of Science)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Geosciences&quot;</td>
<td>&quot;Geosciences, Multidisciplinary&quot;</td>
<td>&quot;Geosciences&quot;</td>
<td>Cornell Dept. of EAS</td>
</tr>
<tr>
<td>5 Geochimica et cosmochimica acta</td>
<td>Paleoceanography</td>
<td>Bulletin of the American Meteorological Society</td>
<td>Geophysical Research Letters</td>
</tr>
<tr>
<td>6 Geology</td>
<td>Journal of Paleontology</td>
<td>American Journal of Science</td>
<td>Geochemistry, geophysics, geosystems</td>
</tr>
<tr>
<td>7 Icarus</td>
<td>Precambrian Research</td>
<td>Earth-science reviews</td>
<td>Journal of Geophysical Research*</td>
</tr>
<tr>
<td>8 Journal of the atmospheric sciences</td>
<td>Geological Society of America Bulletin</td>
<td>Paleoceanography</td>
<td>Annales Geophysicae [Germany]*</td>
</tr>
<tr>
<td>0 Monthly weather review</td>
<td>Geophysical Research Letters</td>
<td>Journal of climate</td>
<td>Geology**</td>
</tr>
<tr>
<td>1 Quaternary science reviews</td>
<td>Landscape Ecology</td>
<td>Journal of Geophysical Research. Space Physics**</td>
<td></td>
</tr>
<tr>
<td>2 Chemical geology</td>
<td>Biogeochemistry</td>
<td></td>
<td>Nature**</td>
</tr>
<tr>
<td>3 Remote sensing of environment</td>
<td>Turkish Journal of Earth Sciences</td>
<td></td>
<td>Science**</td>
</tr>
<tr>
<td>Palaeogeography</td>
<td></td>
<td></td>
<td>Earth and Planetary Science Letters***</td>
</tr>
<tr>
<td>Palaeoclimatology</td>
<td></td>
<td></td>
<td>Field Crop Research***</td>
</tr>
<tr>
<td>Palaeoecology</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Note that Cornell’s EAS faculty seem to cite a lot more physics journals than they do geoscience titles.
Questions?