

# Introduction to Citation Analysis with Sci2

Ista Zahn and Steve Worthington



**The Institute**  
*for* Quantitative Social Science  
at Harvard University

# Download Sci2

- <https://sci2.cns.iu.edu>
- Free registration required
- Download then unzip
- No need to install – just run the sci2 executable from the unzipped sci2 directory

# Workflow Overview

- Obtain citation data to be analyzed
- Load data in Sci2
- Prepare and pre-process data: remove duplicates, extract networks, aggregate etc.
- Analyze: burst detection, geocoding, clustering etc.
- Visualize: Time series, geospatial, network

# Example 1: Temporal Analysis

- Examine trends in publications in the Journal “Scientometrics”
- Main questions:
  - What are the trends in topics studied over time?
  - When did top authors rise to prominence?

# Locate Citations to be Analyzed

- <http://apps.webofknowledge.com.ezp-prod1.hul.harvard.edu/>

The screenshot shows the Web of Knowledge interface in a web browser. The browser's address bar displays the URL: `apps.webofknowledge.com.ezp-prod1.hul.harvard.edu/WOS_GeneralSearch_input.do?product=WOS&SID=3...`. The page header features the "WEB OF KNOWLEDGE<sup>SM</sup>" logo, the tagline "DISCOVERY STARTS HERE", and the "THOMSON RE" logo. Below the header, there is a navigation bar with links: "Go to mobile site", "Sign In", "Marked List (0)", "My EndNote Web", "My ResearcherID", "My Citation Alerts", "My Saved Searches", "Log Out", and "Help". The main content area has tabs for "All Databases", "Select a Database", "Web of Science", and "Additional Resources". Under the "Web of Science" tab, there are sub-tabs: "Search", "Author Search", "Cited Reference Search", "Advanced Search", and "Search History". The "Search" sub-tab is active, displaying the "Web of Science<sup>®</sup>" title and a "Search" section. This section contains three search input fields, each with a dropdown menu for the search type and a magnifying glass icon. The first field contains the text "Scientometrics" and has a dropdown menu set to "Author". Below it, an example search string is provided: "Example: O'Brian C\* OR OBrian C\*", followed by the text "Need help finding papers by an author? Use Author Search." The second field is empty and also has a dropdown menu set to "Author". The third field is empty and has a dropdown menu set to "Publication Name". Each field is preceded by an "AND" button with a dropdown arrow.

Web of Knowledge [v.5.8] - x

apps.webofknowledge.com.ezp-prod1.hul.harvard.edu/WOS\_GeneralSearch\_input.do?product=WOS&SID=3...

WEB OF KNOWLEDGE<sup>SM</sup> | DISCOVERY STARTS HERE | THOMSON RE

Go to mobile site | Sign In | Marked List (0) | My EndNote Web | My ResearcherID | My Citation Alerts | My Saved Searches | Log Out | Help

All Databases | Select a Database | Web of Science | Additional Resources

Search | Author Search | Cited Reference Search | Advanced Search | Search History

Web of Science<sup>®</sup>

Search

Scientometrics in Author

Example: O'Brian C\* OR OBrian C\*  
Need help finding papers by an author? Use Author Search.

AND in Author

Example: O'Brian C\* OR OBrian C\*  
Need help finding papers by an author? Use Author Search.

AND in Publication Name

Example: Cancer\* OR Journal of Cancer Research and Clinical Oncology

# Download Citations

- Download to a delimited plain text format
- WoS only allows 500 records to be exported at a time

The screenshot displays the Web of Knowledge interface. At the top, the browser address bar shows the URL: `apps.webofknowledge.com.ezp-prod1.hul.harvard.edu/summary.do?SID=3BPkoN%40HbOKEIOKpOkO&prod...`. Below the address bar, the page title is "Web of Knowledge [v.5.8]". The main content area shows a search result for "391-411 DOI: 10.1007/S11192-012-0681-1" published in "NOV 2012" with "Times Cited: 0 (from Web of Science)". A purple button labeled "FIND IT @ HARVARD" and a blue link "View abstract" are visible. Below this, the results summary shows "Results: 3,342" and "Page 1 of 335". The "Sort by:" dropdown is set to "Publication Date -- newest to oldest". The "Output Records" section is divided into three steps: Step 1 (Selected Records on page, All records on page, Records 1 to 500), Step 2 (Authors, Title, Source, Full Record, plus Cited References), and Step 3 (Save to: ENDNOTE WEB, ENDNOTE, RefWorks, I Wrote These Publications, Save to other Reference Software, Save). The bottom of the page shows the text "3,342 records matched your query of the 48,507,946 in the data limits you selected." and a footer with "© 2012 Thomson Reuters" and links to "Terms of Use", "Privacy Policy", and "feedback".

Web of Knowledge [v.5.8]

apps.webofknowledge.com.ezp-prod1.hul.harvard.edu/summary.do?SID=3BPkoN%40HbOKEIOKpOkO&prod...

391-411 DOI: 10.1007/S11192-012-0681-1 Published: NOV 2012  
Times Cited: 0 (from Web of Science)

FIND IT @ HARVARD [View abstract]

Results: 3,342 Show 10 per page Page 1 of 335 Go Sort by: Publication Date -- newest to oldest

**Output Records**

**Step 1:**

- ☐ Selected Records on page
- ☐ All records on page
- ☒ Records 1 to 500

**Step 2:**

- ☐ Authors, Title, Source
  - ☐ plus Abstract
- ☒ Full Record
  - ☒ plus Cited References

**Step 3:** [How do I export to bibliographic management soft

Save to: ENDNOTE WEB ENDNOTE RefWorks I Wrote These Publications Save to other Reference Software Save

3,342 records matched your query of the 48,507,946 in the data limits you selected.

View in: 简体中文 English 日本語

© 2012 Thomson Reuters Terms of Use Privacy Policy Please give us your feedback on using Web of Knowledge.

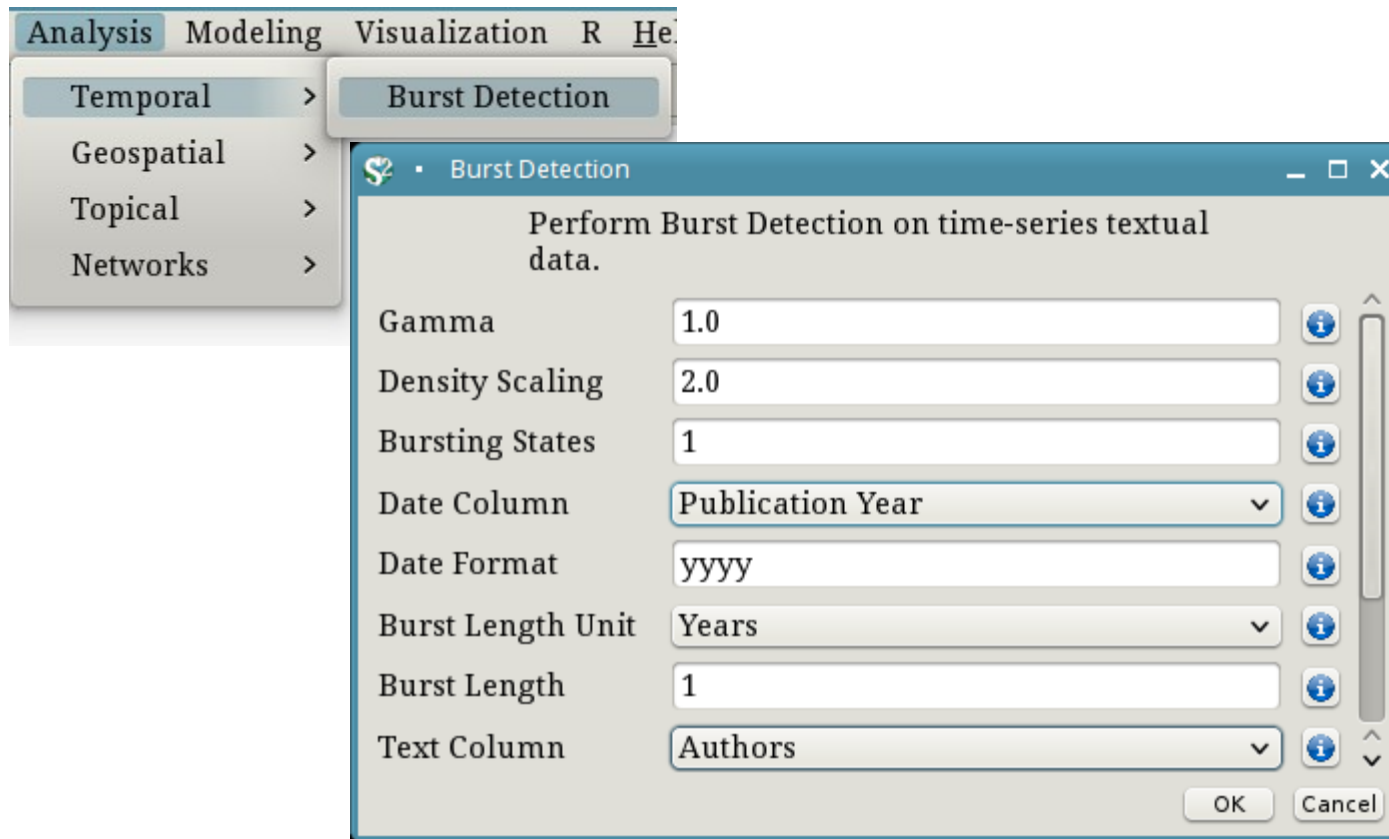
# Rename and Load Saved Records

- Rename the save records from “filename.txt” to “filename.isi”
- If you have multiple record files, merge them
- Load records in Sci2

- Check console log 

# Temporal Analysis: Author Longevity

- Burst detection for authors





# Temporal Analysis: Author Prominence

- Clean up burst detection results



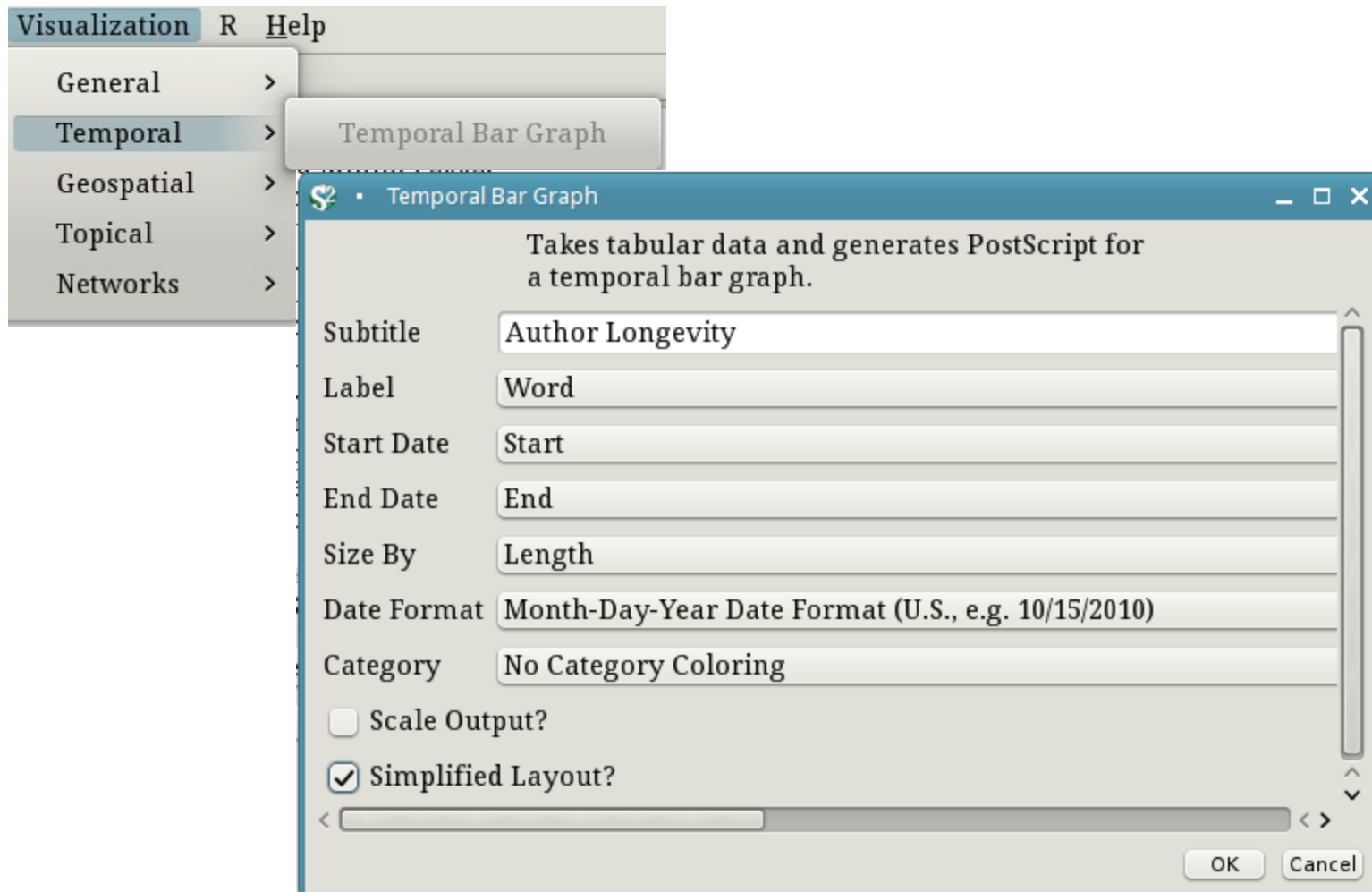
- Replace blank End date with “2012”

Word	Level	Weight	Length	Start	End
Daniel, HD	1	4.0709429897	8	2005	2012
Park, HW	1	6.0418679225	2	2011	
FRAME, JD	1	4.6370688996	8	1980	1987

- Save and re-load into Sci2

# Temporal Analysis: Author Longevity

- Visualize author longevity



# Temporal Analysis: Author Longevity

- Visualize author longevity

