Introduction to SciVal

Basic Structure and Usage of the SciVal Database
Workshop Structure

- **Intro to SciVal**
  - What is SciVal/How does it work?
  - 3 practice use cases:
    - Assessing BU research performance vs other institutions
    - Identifying potential collaborators
    - Assessing BU research performance in a custom topic area
  - Q&A and practice time
What is SciVal

- SciVal is a bibliometric platform powered by the Elsevier Scopus database
- Provides data built on publications as the primary unit. Everything is built up from publications
- Data can provide insight from individual publication up to global trends
Database Source/Structure

- All SciVal data and visualizations are based on Scopus publication data
- Publications are updated in Scopus daily
- SciVal updates from Scopus weekly
- For most metrics SciVal maintains a 5 year window, some allow for larger windows
Cautions When Using SciVal

- SciVal contains all publication types: articles, books, conference papers, chapters, and more. SciVal is strongest in areas where traditional journal publications are the standard publication in the field.

- SciVal is functional, though not as robust, in fields where other output is more common, for example:
  - Computer Science: Greater focus on conference proceedings, frequently without a DOI and thus missed in the database.
  - Economics: Fewer authors per paper, greater focus on books, makes comparative use tricky.

- Field is typically dictated by the discipline of the journal, not the content of the article.
Database Source/Structure: Keywords

- Large portions of the SciVal database and its search functions are based on keywords
- Papers are classified through a bigram analysis to flag keyword pairs in addition to single keywords
- Most custom searches in SciVal create groups of publications around a keyword, refine those groups, and then allow for analysis of the resulting publication set
Database Source/Structure: Topics + Clusters

- SciVal includes pre-made topics and topic clusters
- These are generated through the combination of the bigram word analysis and citation analysis
- This generated groups (by keyword) of papers that heavily cite each other, but don’t cite much outside the group (Topics)
- The same analysis found the same pattern at a higher level. Clusters of topics cite each other much more frequently (Topic Clusters)
- The bigram is then used to uniquely identify/name the topic or topic cluster
Basic Use: Overview

- Good for high level metrics of a whole institution, researcher, or topic
- Allows creation of a variety of visuals or tables
- Provides a starting point to explore data before crafting more specific visuals
- Also contains information about media mentions and patents, though this information is limited in some cases
Basic Use: Overview

Selection of institutions, researchers, and/or topics

High level metrics for the selected entity

Tabs to access other SciVal functions
Basic Use: Benchmarking

- Benchmarking offers a streamlined way to create visuals and data tables
- Good for comparison of institutions, researchers, or topics over time (Benchmarking has the greatest flexibility in year range of any portion of SciVal)
- Flexible in that different types of entity (i.e. researchers and institutions) can be put on the same visualization
Basic Use: Benchmarking

Selection of entities to include on the graphic, can mix and match types

List of currently displayed entities, the color can be changed by clicking on the symbols here

Selection of display variables, 2 by default, some horizontal axis selections allow for the selection of 3
Example Case: Comparing Institutions in Benchmarking

- How does BU compare in publication quantity/quality to Boston College, Northeastern, Harvard, and MIT in Biophysics over the last 5 years?
  - Skills: Define schools, Benchmarking tab, Selecting category within ASJC, Changing metrics in a visual
Basic Use: Collaborations

- Collaborations is a tool to quantify both current and potential collaborators at the researcher and institutional level
- Provides both list and visual output
Basic Use: Collaborations

Entity selection for collaboration, can be an institution, person, or country/region

Map displaying collaborations, can be zoomed in to see greater detail on a region/country
Example Case: Finding Collaborators

- An education grant requires collaboration with a European institution, who are possible partners that have collaborated with BU in the past?
  - Skills: Collaborations tab, ID co-authors, ID institutions with existing collaboration, ID with a certain discipline
Basic Use: Trends

- Provides some similar and some more in-depth information on topics, publication sets, and research areas
- Can give information at all levels
- Has some trend/momentum metrics that can indicate the direction of work in a field
Basic Use: Trends

Entity must be a publication set, topic, or topic cluster

Focus on metrics over time to see “trends” in the field

Word cloud for the entity showing both frequency over time (color) and relevance (size)
Basic Use: Reporting

- Allows the creation of printouts containing the data tables and visualizations from other portions of SciVal
- Essentially anything from the other tabs can be added to a report
- Reports can then be labeled and printed
- Useful, albeit limited function
Defining Custom Groups

- SciVal Allows you to define custom groups of researchers or publications
  - Researchers manually searched and added to a group
  - Publications use keywords and other filters to create a list
- Small groups (under 5,000) publications can be generated immediately in most cases
- Large groups (above 5,000) will be processed as soon as possible, usually within 2 or 3 days. No processing is done on the weekends
Defining Custom Groups

Large range of filtering options to refine custom group as needed

Any filter selection can be used as a limit or exclude function

List of publication counts in total and counts that are within a filter group
Example Case: Analyzing Output in a Custom Research Area

- Where does BU rank domestically in FWCI for publications related to concussions?
  - Skills: Define a custom research area, Trends tab, Analyzing top 100 in more detail, Filtering data for visuals