LEXICON

• ESI: Essential Science Indicators
• JCR: Journal Citation Reports
• JIF: Journal Impact Factor
• WoS: Web of Science

This document was compiled by Thomson Reuters® for InCites, ESI and JCR customers attending training courses. This document or its copies are not for public use or publication.

Information provided was valid at the time of release, please refer to the version date above. If you identify errors, please contact your trainer.
The existing version of InCites can be accessed from [http://incites.isiknowledge.com](http://incites.isiknowledge.com)

**Getting more help**

- Training videos on InCites: [http://wokinfo.com/training_support/training/incites/](http://wokinfo.com/training_support/training/incites/)
Course plan (IC2BA02)

1. Prerequisite
   1. InCites quick tour: http://youtu.be/Q2Idsq1rT4g (5’ videos)
   2. Download training material: http://bit.ly/1gZ4JfDh
      - InCites Quick reference guide
      - InCites metrics handbook

2. InCites: Benchmarking & Analytics
   1. Thomson Reuters’ Research Analytics
   2. InCites data and metrics
   3. Using InCites for evaluation and benchmarking

3. Work toward an example (optional)
4. Give us some feedback about today’s session
   bit.ly/TrainingWoS

Summary

This Course is designed to give you an overview of the Journal and Highly Cited Data on Thomson Reuters’ InCites platform. We will cover the data and the key features available in InCites Benchmarking & Analytics (B&A).

Prior to attending this session, we ask you to watch the quick tour videos that will give you a general overview on these services.

• InCites B&A quick tour: http://youtu.be/Q2Idsq1rT4g

• Training material is available from: http://bit.ly/1gZ4JfDh

  • ESI quick tour: http://youtu.be/7xxKl3Oj_QI
  • JCR quick tour: http://youtu.be/mgUeK0rRJgM

This session will be divided in 3 parts:

• Thomson Reuters’ Research Analytics
• InCites data and metrics
• Using InCites for evaluation and benchmarking
2.1.1 The basics of Research Analytics

• No data for the sake of it, what is your question?
• What do you want to know?
• What success looks like?

No data for the sake of it, what is your question?
Data is used as evidence to support a policy or a strategy, from development to implementation and review.
Data must have a set of features to be acted upon:
  • Trusted source of data and calculation methods,
  • Robustness to slight changes in environment and outliers,
  • Reflect observed notable changes.

What do you want to know?
Prior to any evaluation or analysis, the key problem should be explicit and formulated in one or more questions:
  • What are my most relevant publications for the last 5 years?
  • Is University X an important partner?

What success looks like?
Indicators are used to measure how an entity or a groups of entities perform.
Measuring success is linked to initial hypothesis about the measured process, hence determining selected indicators:
  • Relevant publications are highly cited by peers in the discipline, high normalised citation score.
  • Important partner universities collaborate internationally notably with China, organisations with high % on international collaborations co-authoring with China.
Indicators are necessarily limited in the scope they cover. Despite the flurry of new research performance metrics published every week, they rarely measure more than limited aspects of research. Including metrics such as the H index.

To cover all aspects of the evaluated activity needed to answer the initial question. Generally, that means all or part of the followings aspects:

- Productivity – how many papers
- Academic impact – absolute of relative citation counts,
- Collaborations – with who, where are they working,
- Specialisation – what disciplines.
2.2.1 InCites publication data

- A single data source for publications: Web of Science
- Publication window = citation window,
- Baselines calculated for all documents and the same documents,
  * late 2014

InCites publication and citation data (included records)

- All journals, proceedings and books indexes from the Web of Science Core collection.
- Selected records: initially published in the last 10 years (eg 2004-2013), but this timeframe is not a constant window. Later version of InCites will include earlier publications and bi-monthly updates of the latest publications.
- Citations are calculated from the same selected source records (included indexes and timeframe).
- Baselines are calculated from the same publication and citation window.

InCites vs. Web of Science

Publication and citation data are extracted from the Web of Science a few weeks before an InCites update is released. There is not “live” stream of data between the two platforms. When comparing number of publications and citations between the platforms, some differences can be observed:

- publication and citation indexes and time windows need to be similar,
- records in Web of Science can change between the InCites extract and the latest Web of Science update (made weekly),
- fields such as the “Organisation Enhanced” in Web of Science are updated at a different interval.
### 2.2.2 InCites other data

<table>
<thead>
<tr>
<th>Recorded Future</th>
<th>Beyond publication data</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Scan public web sources, including news publications, “high-caliber” blogs,</td>
<td>• patents</td>
</tr>
<tr>
<td>social media platforms, financial databases, government websites, and much</td>
<td>• grant data (mainly US and UK information),</td>
</tr>
<tr>
<td>more.</td>
<td>• Planned for late 2014.</td>
</tr>
<tr>
<td>• Sources are analysed through text search, data visualization, natural</td>
<td></td>
</tr>
<tr>
<td>language processing, and entity extraction.</td>
<td></td>
</tr>
</tbody>
</table>

Data currently available

• Recorded Future is a start up company created in 2009
  ([https://www.recordedfuture.com/about/](https://www.recordedfuture.com/about/)).

Data included later

• Patent data will come from Thomson Reuters Derwent Innovation Index.
• Grant data are sourced from Thomson Reuters information systems.
2.2.3 Coverage

- Web of Science is a multidisciplinary index of natural and social sciences as well as the arts and humanities,

- However, in some disciplines research outputs are better covered through papers, and books over "non traditional" outputs such as creative work, live performances, recordings or exhibitions.

Example of coverage: Documents produced in Australia by FoR codes (InCites)

<table>
<thead>
<tr>
<th>FoR</th>
<th>Documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD multidisciplinary</td>
<td>20,785</td>
</tr>
<tr>
<td>18 law and legal studies</td>
<td>2,020</td>
</tr>
<tr>
<td>19 studies in creative arts and writing</td>
<td>2,671</td>
</tr>
<tr>
<td>12 built environment and design</td>
<td>3,062</td>
</tr>
<tr>
<td>22 philosophy and religious studies</td>
<td>4,792</td>
</tr>
<tr>
<td>21 history and archaeology</td>
<td>6,407</td>
</tr>
<tr>
<td>10 technology</td>
<td>6,566</td>
</tr>
<tr>
<td>20 language, communication and culture</td>
<td>6,849</td>
</tr>
<tr>
<td>14 economics</td>
<td>7,980</td>
</tr>
<tr>
<td>13 education</td>
<td>9,332</td>
</tr>
<tr>
<td>15 commerce, management, tourism and services</td>
<td>10,662</td>
</tr>
<tr>
<td>04 earth sciences</td>
<td>13,902</td>
</tr>
<tr>
<td>01 mathematical sciences</td>
<td>13,993</td>
</tr>
<tr>
<td>08 information and computing sciences</td>
<td>14,028</td>
</tr>
<tr>
<td>05 environmental sciences papers</td>
<td>14,150</td>
</tr>
<tr>
<td>07 agricultural and veterinary sciences</td>
<td>20,066</td>
</tr>
<tr>
<td>16 studies in human society</td>
<td>20,693</td>
</tr>
<tr>
<td>02 physical sciences</td>
<td>24,166</td>
</tr>
<tr>
<td>03 chemical sciences</td>
<td>25,746</td>
</tr>
<tr>
<td>17 psychology and cognitive sciences</td>
<td>34,418</td>
</tr>
<tr>
<td>09 engineering</td>
<td>47,014</td>
</tr>
<tr>
<td>06 biological sciences</td>
<td>52,579</td>
</tr>
<tr>
<td>11 medical and health sciences</td>
<td>149,493</td>
</tr>
</tbody>
</table>

Example of coverage: Documents produced in Australia by FoR codes (source: InCites) and the type of outputs for some of them (source: ERA 2010 data).
The basic element of InCites data is a Web of Science document or publication. These documents are defined by three types of information:

**Bibliographic information**
- Author names,
- Publication year (cover date),

**Metadata**
- Author affiliation: organisation, country,
- Author identification such as ResearcherID codes,
- Research Areas: the research category of a publication can be attributed through different mechanisms. For instance, the web of science areas are attributed to the outlets such as journals,
- Document type: each publication can be given one or more type. Documents classified as articles or reviews are generally called “paper”,
- Open Access status.

**Citation network**
- References cited in each document,
- Documents citing the source document.
2.2.5 Productivity metrics: measuring outputs

Number of documents
- In InCites Dataset > In the Research Areas scheme > Selected filtering,
- All documents of specific types can be selected: e.g. papers (articles, reviews, proceedings papers and book chapters),
- Period analysis or trend analysis (annual evolution).

Number of documents

All records from the Web of Science core collection are part of the InCites data but the total number of documents given in InCites will vary upon the selected filtering notably:
- The Research Areas scheme: under each scheme, the total “pool” of documents can vary (e.g. The Essential Science Indicator scheme only includes papers published in journal with a ESI category),
- The document type: all documents types across the WoS Core collection are available and the specific types can be selected. Typically, are called papers documents classified as articles, reviews, proceedings papers and book chapters. Documents in the Web of Science can have more than one type, but in InCites each document will be attributed a single type (e.g. A book chapter published as an article in a journal, might have an article + book chapter type, but will be considered as an article in InCites). The Web of Science scheme will include all the documents available.

Trends

The evolution of the number of documents can be analysed over a period (time frame) or through the annual variation (trend graph) within this time frame.
2.2.6 Impact metrics: influence and comparison

Citation Impact
Measure influence.
- Times cites,
- Citation Impact,
- % documents cited.

Citation benchmarks
Measure impact relative to peers.
- Normalized Citation Impact,
- Average percentile,
- Percentage of top papers,
- Percentage of highly cites papers.

1: Global baseline = 1*
2: NCI=1.45 > 1, above peers
3: NCI=1.02 = 1, average
* Calculated value not exactly at 1.00

Citation impact
The number of citations accrued by publications is an indicator of their overall influence. The higher the value and rate of citation the higher the influence is.

Total cites: raw number of citations, it measures influence of papers, authors, institutions,
Citation impact: average number of citations per document (often referred as CPP),
Percentage of cited documents: can also be seen as the percentage of uncited documents.

Citation benchmark
A given publication will attract a variable level of cites depending on a series of factors. The observed cites will therefore be a function of the age of the publication and these factors as summarised below:

Cites to publication=f(influence of the publication, age, document type, field, other)

"other" residual factors can notably include the number of authors and affiliations, the geographic origin of the publication.

To take into account the effect of “age, document type, field” InCites provides a series of normalised indicators that enable the comparison of a pair of publications or a pair of document sets.

Normalized Citation Impact: the average number of citation to the selected publications normalised for the publication year, the category and the document type.

Average percentile: the average rank of papers in the selected publications that are ranked by number of Cites each year and in individual field. Papers in the top 3% will have be ranked in the 97th percentile.

Percentage of top papers: The percentage of papers in the selected publications that are ranked in the top 1% or 10% by number of Cites each year and in individual field.

Percentage of highly cites papers: as above but it identify the top 1% papers in individual ESI categories over the latest ESI period (last 10 years).
2.3 Demonstration

InCites Benchmarking & Analytics course (IC2BA02)
THANK YOU

More questions?
jfdh@thomsonreuters.com

GIVE US SOME FEEDBACK:
BIT.LY/TRAININGWOS