Research Intelligence

Research Metrics Guidebook

Table of contents

ELSEVIER

ELSEVIER

1.0 Introduction: Elsevier's approach to research metrics

The quote often attributed to Albert Einstein, but perhaps more properly attributed to William Bruce Cameron[1], is often referred to when writing a foreword such as this:

"Not everything that counts can be counted, and not everything that can be counted counts".

This is undoubtedly true, but does not follow that nothing should be measured. There is much that can be counted that is important and provides valuable perspectives on trends in academia, and there is an increasing emphasis on this in the world of research today.

The field of metrics relies on specialized academic study for its development and to capitalize on advances in technology; these scholarly outputs are increasingly being used by people involved in research, whether directly by conducting research or indirectly in a supporting or enabling role. The majority of these people would not consider themselves to be experts in metrics or their use: this Guidebook is aimed at those people, and at supporting the use of metrics in an appropriate and responsible way.

Many of the metrics in this Guidebook were developed with the intent of identifying and understanding research impact. The comprehensive suite of research metrics embedded throughout Elsevier's Research Intelligence portfolio is designed to help facilitate evaluation and provide a better view of research outcomes.

Research metrics aim to give a balanced, multi-dimensional view for assessing published research. Built on the depth and breadth of its data sources, Elsevier works with researchers, publishers, bibliometricians, librarians, institutional leaders and others in academia to offer an evolving basket of metrics that complement qualitative insights. Throughout our solutions, you can access multiple metrics, including at the journal, article, author and institutional levels.

This Guidebook is intended to be a straightforward, practical companion to the use of tools like Scopus and SciVal, which are a part of the Research Intelligence portfolio of solutions. It provides some facts about how the data underlying the metrics are used, how the metrics are calculated and displayed, and about variables besides performance that can affect the metrics. It also provides some suggestions about situations where the metrics are useful, when care should be taken, and how shortcomings may be addressed.

We provide only two rules in the use of metrics within this Guidebook: always use more than 1 metric to give insights into a question, and always use metrics in conjunction with peer review and/or expert opinion to support any conclusions. This "triangulation" of approaches will increase the reliability of conclusions drawn if these distinct types of information reinforce each other, and will flag areas for further investigation if they do not. Beyond this, there are no black-and-white rules. The best approach is to use common sense to ensure metrics are used responsibly.

1.1 Journal metrics

Journal level metrics continue to be an important part of the basket of metrics, complementing new and alternative metrics to provide a multi-faceted view of a journal's impact. On Scopus, you will find an evolving and expanding suite of journal metrics that go beyond just journals to include most serial titles, including supplements, special issues and conference proceedings. Freely available on Scopus you will find CiteScore metrics, SCImago Journal Rank (SJR) and Source Normalized Impact per Paper (SNIP).

1.2 Article-level metrics

PlumX Metrics is now the primary source of article-level metrics in the Research Intelligence portfolio alongside the Scopus Citation Count (including percentile benchmarking) and Field-Weighted Citation Impact. PlumX Metrics provide insights into the ways people interact with individual pieces of research output (articles, conference proceedings, book chapters, and many more) in the online environment.

Examples include when research is mentioned in the news or is tweeted about. Collectively known as PlumX Metrics, these metrics are divided into five categories- Usage, Captures, Mentions, Social Media, and Citations- to help make sense of the huge number of metrics available and to enable analysis through comparing 'like with like'. PlumX gathers and brings together appropriate research metrics for all types of scholarly research output.

1.3 Author and institutional metrics

Author and institutional metrics can help you assess an entity's research output and scholarly impact. The depth and breadth of content on Scopus, combined with industry-leading technology powering algorithmic and systematic author and institutional entity disambiguation, provides the quality data needed to build accurate measurements of impact.

With Scopus you can easily analyze and track an individual or institution's publication history. In addition to finding total citation and document counts, here are a few examples of metrics and tools available:

- *h*-index and *h*-graph: Rates a scientist's performance based on his or her career publications, as measured by the lifetime number of citations each article receives. The measurement depends on both quantity (number of publications) and quality (number of citations) of an academic's publications.

- Citation overview tracker: An adjustable table that includes the number of times each document has been cited per publication year.

- Analyze author output: A collection of in-depth and visual analysis tools designed to provide a better picture of an individual's publication history and influence.

1   http://quoteinvestigator.com/2010/05/26/everything-counts-einstein/
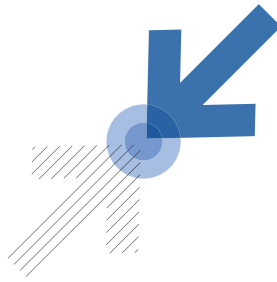
ELSEVIER

2.0 Scopus: the primary data source for Elsevier's Research Metrics

Many metrics and information displayed in Elsevier's Research Intelligence solutions at the time of writing this Guidebook are based on Scopus. This section highlights some of the key details of Scopus that are useful in understanding the metrics, particularly those used in SciVal. The reader can find more extensive information about Scopus online[2].

2.1 Scopus content

The independent and international Scopus Content Selection and Advisory Board reviews titles for inclusion in Scopus on a continuous basis. Information about the process and the acceptance criteria is available online[3]. The Board considers journals, conference proceedings, trade publications, book series, and stand-alone books for inclusion. Scopus indexes more than 70 million publications. Reference lists are captured for the 65 million records published from 1970 onwards. The additional 6+million pre-1970 records reach as far back as the publication year 1788.

2.1.1 Scopus content and SciVal

While Scopus covers content going back to 1788, SciVal uses Scopus content from 1996 onwards so that the Citation Counts displayed in SciVal are based on uninterrupted years of data and reflecting most current trends.

2.1.2 Books and SciVal

Scopus indexes both book series, and stand-alone books. "Books" in SciVal refers to stand-alone books only; their characteristics, and the fact that they do not have journal metrics, sets them apart from the series of books, journals, conference proceedings and trade publications. Scopus links citations to the individual chapters of edited volumes when the information provided by the author allows this, and otherwise to the edited volume itself. Scopus makes either one link or the other, but not both. Our solutions credit a researcher who is the author of a book chapter with the count of citations linked to that particular chapter; it credits a researcher who is the editor of the entire volume with the citations linked to the volume plus those linked to all the individual chapters to ensure that the editor is credited with the full citation impact of their scholarly contribution.

2.2 Data currency in Scopus and SciVal

The data in Scopus are updated daily. The data in SciVal are updated every week. SciVal takes a copy of the Scopus database that is then structured to optimally support its metrics and functionality. This means that SciVal may be slightly behind Scopus in its data currency.

2.3 Author Profiles

2.3.1 Author Profiles in Scopus

Scopus is the only database in the world which has invested in automatically grouping the publications it indexes into those published by a single author. Author identifiers (Author Profiles) group together publications belonging to one author, and they have 2 modes of input:

Publications are automatically grouped into Author Profiles using a matching algorithm:

- This algorithm looks for similarities in author name, as well as affiliation, journal portfolio, and discipline to match publications together. Users may notice that multiple name variants are grouped within one Author Profile, which indicates the value of this algorithm

- The information provided by authors is not always consistent or complete, and even if this were the case, the mobility of authors means that there is always some doubt about whether some publications belong together. In situations like these, a balance needs to be made between the precision, or accuracy, of matching, and the recall, or completeness, of the groups formed, and increasing one will reduce the other

- The Scopus algorithm favors accuracy, and only groups together publications when the confidence level that they belong together, the precision of matching, is at least 99%, such that in a group of 100 papers, 99 will be correctly assigned. This level of accuracy results in a recall of 95% across the database: if an author has published 100 papers, on average 95 of them will be grouped together by Scopus

- These precision and recall figures are accurate across the entire Scopus database. There are situations where the concentration of similar names increases the fragmentation of publications between author profiles, such as in the well-known example of Chinese authors. Equally there are instances where a high level of distinction in names results in a lower level of fragmentation, such as in Western countries

- A publication that has multiple co-authors will be part of multiple Author Profiles

Publications are manually reassigned based on feedback.

The matching algorithm can never be 100% correct because the data it is using to make the assignments are not 100% complete or consistent. The algorithm is therefore supplemented by feedback received from the sector, including from the Open Researcher and Contributor ID (ORCID) initiative[4], and that feedback is used to enhance the profiling of authors by the Scopus Author Feedback Team[5].

2.3.2 Authors and Researchers in SciVal

The presence of Author Profiles in Scopus enables enormous flexibility for SciVal users. Every Author Profile is available to SciVal users to define and view any Researcher in the world in real time, whether they are based at the user's institution, are a collaboration partner in another country, or they are a Researcher who was unknown until today. Users can build on the Author Profiles to define as many Researchers as they like in their SciVal accounts, and subsequently use their Researchers as the basis to create an unlimited number of:

- **Groups of Researchers.** These could represent actual research teams, departments or other organizational groups, international collaboration networks, or models of research teams that are being considered

2 http://www.elsevier.com/online-tools/scopus
3 http://www.elsevier.com/online-tools/scopus/content-overview
4 http://orcid.org/
5 http://www.scopusfeedback.com/
6 http://www.ref.ac.uk/

- **Publication Sets.** These could be pieces of a researcher's output, such as publications from a particular affiliation, or those funded by a particular award

- **Groups of Publication Sets.** These could represent a selection of publications being submitted to a national evaluation, such as the Research Excellence Framework in the United Kingdom[6], for example

SciVal distinguishes between its use of the terms "authors" and "Researchers":

- **Authors are the automatically created Scopus Author Profiles.** The author count in the Overview module, for example, is a count of unique Author Profiles; this may be an over-estimation of the number of researchers, because the recall rate of 95% means that a researcher's publication portfolio might be split over multiple author profiles

- **Researchers are entities which have been created with the benefit of human input.** SciVal users can combine Author Profiles, remove any publications that do not belong to the author of interest, and search for individual publications that should be added in order to create a Researcher. The Profile Refinement Service that populates SciVal with Researchers and Groups of Researchers on behalf of an institution is a similar manual process, performed by Elsevier on behalf of a customer. SciVal distinguishes these manually created Researchers, whether created by users or by Elsevier, from the automatically generated Author Profiles. All of these manual enhancements are fed back to Scopus and used to improve the quality of the source database for all users of Scopus data. This means that once the feedback has been processed by Scopus, each Researcher in SciVal is represented by a single Author Profile which can be automatically updated as new publications are indexed

2.4 Affiliation Profiles

2.4.1 Affiliation Profiles in Scopus

Scopus is the only database in the world which has invested in automatically grouping the publications it indexes into those published by a single affiliation. These groups of publications belonging to one affiliation are called Affiliation Profiles, and they have 2 modes of input:

Publications are automatically grouped into Affiliation Profiles using a matching algorithm:

- This algorithm looks for similarities in affiliation name, as well as addresses, to match publications together. Users may notice that multiple name variants are grouped within one Affiliation Profile, which indicates the value of this algorithm. Scopus makes use of an authoritative database that contains over 70,000 manually verified institutional name variants to match publications together

- The information provided by authors is not always consistent or complete, so that there is always some doubt about whether some publications belong together; in situations like these, a balance needs to be made between the precision, or accuracy, of matching, and the recall, or completeness, of the groups formed, and increasing one

will reduce the other

- The Scopus algorithm favors accuracy, and only groups together publications when the confidence level that they belong together, the precision of matching, is at least 99%, such that in a group of 100 papers, 99 will be correctly assigned. This results in a recall of 93% across the database, such that if an affiliation has published 100 papers, on average 93 of them will be grouped together by Scopus, and the others will be in 1 or more separate groups

- A publication that has co-authors with multiple affiliations will be part of multiple Affiliation Profiles

Publications are manually reassigned based on feedback. The matching algorithm can never be 100% correct because the data it is using to make the assignments are not 100% complete or consistent. The algorithm is therefore supplemented by feedback received from the official authority of the affiliation in question

2.4.2 Affiliations and Institutions in SciVal

The presence of Affiliation Profiles in Scopus brings enormous benefits for SciVal users, with metrics being pre-calculated and available to view at the click of a mouse. SciVal users also benefit from the availability of Groups of Institutions, such as those found in each of the American states.

SciVal distinguishes between its use of the terms "affiliations" and "Institutions":

- Affiliations are the automatically created Scopus Affiliation Profiles. A medical school will be a separate Affiliation Profile from a university, for instance

- Institutions are groupings of related Affiliation Profiles which have been manually created as a convenient starting point for SciVal users; more than 10,000 Institutions have been pre-defined and are available in SciVal. Any medical schools are always grouped together with the university in the SciVal Institutions

2.5 Do publications belong to Institutions or to Researchers?

Researchers are mobile, and tend to change affiliations during their careers. This leads to 2 perspectives as to whether publications "belong" to Institutions or to Researchers:

- **The "Institutional perspective"** is typically that publications should remain assigned to them even when the researchers that authored these publications have moved. In other words, the publications are not mobile, despite their authors moving around

- **The "Researcher perspective"** is generally that publications should be just as mobile as their authors, and should move from affiliation to affiliation as their authors' careers develop

These approaches are both needed to completely support questions that are asked in different situations. SciVal offers both of these perspectives, because publications are linked to both Affiliation and

Author Profiles independently of each other:

- Institutions and Groups of Institutions in SciVal take the "Institutional perspective"

- Researchers, Groups of Researchers, Publications Sets in SciVal take the "Researcher perspective"

2.6 Organization-types

2.6.1 Organization-types in Scopus

Organization-types are assigned to Scopus Affiliation Profiles based on their primary functions. This function is often very clear from the name of the affiliation, and the organization's website is checked for guidance if there is any doubt.

Scopus assigns affiliations to the following organization-types: university, college, medical school, hospital, research institute, corporate, law firm, government, military organization, and non-governmental organization.

2.6.2 Organization-types in SciVal

The organization-types used in SciVal are based on aggregations of the Scopus organization-types to group similar functions together, and to simplify the options for the user. SciVal uses 5 organization-types: Academic, Corporate, Government, Medical, and Other. These are composed of the following Scopus organization-types:

- **Academic:** university, college, medical school, and research institute

- **Corporate:** corporate and law firm

- **Government:** government and military organization

- **Medical:** hospital

- **Other:** non-governmental organization

2.7 Journal metrics in Scopus and SciVal

Scopus and SciVal use several journal metrics that have been developed by academic research teams, and whose methodology has been published in peer-reviewed journals. Information about these metrics is available online, and all the metrics values are also available for free[7].

These metrics are:

- **CiteScore metrics[8]. CiteScore** calculates the average number of citations received in a calendar year by all items published in that journal in the preceding three years. The calendar year to which a serial title's issues are assigned is determined by their cover dates and not the dates that the documents were made available online first. CiteScore metrics are part of an evolving basket of metrics that will continue to grow with input and guidance from the research community.

- **SNIP[9]. Source-Normalized Impact per Paper** is a ratio between the "Raw Impact per Paper", a type of Citations per Publication calculation, actually received by the journal, compared to the "Citation Potential", or expected Citations per Publication, of that journal's field. SNIP takes differences in disciplinary characteristics into

account, and can be used to compare journals in different fields. The average SNIP value for all journals in Scopus is 1.000

- **SJR[10]. SCImago Journal Rank** is a prestige metric, whose methodology is similar to that of Google PageRank. It weights the value of a citation depending on the field, quality and reputation of the journal that the citation comes from, so that "all citations are not equal". SJR also takes differences in the behavior of academics in different disciplines into account, and can be used to compare journals in different fields. The average SJR value for all journals in Scopus is 1.000

Journal metrics, apart from CiteScore, are not calculated for trade publications.

2.8 Other sources for metrics in Research Intelligence solutions

Newsflo

Newsflo offers researchers and academic institutions a way to measure the wider impact of their work by tracking and analyzing media coverage of their publications and findings. Demonstrating the impact of research on society is likely to become more and more important with increasing competition for research funding and attracting students. Newsflo's innovative service will therefore provide input for the growing recognition of alternative metrics as a means to demonstrate impact, supplementing the more traditional impact measures based on citations. Pioneered by Newsflo Ltd in 2012, the start-up was acquired by Elsevier in 2015 and is currently integrating its media monitoring solutions into the Elsevier ecosystem.

PlumX Metrics

PlumX Metrics provide insights into the ways people interact with individual pieces of research output (articles, conference proceedings, book chapters, and many more) in the online environment. Examples include, when research is mentioned in the news or is tweeted about. Collectively known as PlumX Metrics, these metrics are divided into five categories- Usage, Captures, Mentions, Social Media, and Citations- to help make sense of the huge number of metrics available and to enable analysis through comparing 'like with like'. PlumX gathers and brings together appropriate research metrics for all types of scholarly research output.

Patents

Patent-related metrics serve as additional tools to demonstrate research impact. Our solutions like SciVal identify and count citations which research papers have received from patents. From the perspective of a research publication, these would be "forward citations" indicating whether the research results have subsequently been used in the patent world. It is important to remember that patents are published and can only become available for use in research metrics around 18 months after the application date. We look at five of the largest patent offices: EPO (European Patent Office), USPTO (US Patent Office), UK IPO (UK Intellectual Property Office), JPO (Japan Patent Office) and WIPO (World Intellectual Property Organization).

---

7 https://www.scopus.com/sources

8 https://www.elsevier.com/__data/assets/pdf_file/0007/652552/CiteScore-metrics-The-Basics.pdf

9 http://www.sciencedirect.com/science/article/pii/S1751157710000039

10 http://www.sciencedirect.com/science/article/pii/S1751157710000246

ELSEVIER

3.0 Selection of appropriate metrics

The ideal situation, when making research management decisions, is to have 3 types of input: peer review, expert opinion, and information from a quantitative evidence-base.

When these complementary approaches "triangulate" to give similar messages, the user can have a higher degree of confidence that their decision is robust. Conflicting messages are a useful alert that further investigation is probably a good use of time.

It is also preferable that an evidence-base is used to illuminate a question from various angles. Multiple people are often asked to give their professional judgment about a question, and more than 1 peer review is typically sought; in just the same way, triangulating information about the same question from an evidence-base by using 2, 3 or even more different metrics will also ensure that the insights gained in this "corner of the triangle" are the most reliable they can be.

There are not really any strict rules about the selection of which metrics to use, besides approaching a question from more than 1 direction. The most appropriate metrics will always depend on the particular question that the user is asking. The best approach is to highlight some key points that are important to keep in mind, and for the user to apply their common sense.

SciVal offers a broad range of metrics to enable triangulation from the evidence-base, and to cater for the enormous variety of questions that users will ask. It is a rich and powerful resource of information, and can be used responsibly and appropriately by keeping a few facts in mind, as a complement to other sources of information. These facts are the focus of this section.

3.1 Clarity on the question being asked

The aim of using data and metrics as input into decision making is that any differences observed should reflect differences in performance. This will be the case if the user selects metrics that are suitable to answer their question, which in turn relies on 2 important factors:

- The question that is being asked is clearly articulated

- The user is aware of other factors, beyond performance, that can influence the value of a metric. These may or may not be important in the context of the questions being asked, but this judgment can only be made once that question has been clearly articulated

The types of questions asked typically fall into 3 groups:

- **Evaluation of performance**, such as is conducted by a national body on its research institutions for the purposes of allocating national funding, or by a line manager to provide input into career development discussions. It is typically very important in these situations that variables besides differences in performance have been accounted for to ensure that the assessment is fair; it would not be advisable to compare chemistry and immunology using metrics that do not take into account the tendency for higher output and citation rates in immunology, for instance

- **Demonstration of excellence**, such as that which may support an application for competitive funding, or that which may be used for promotional purposes to attract post-graduate students to a research institution. The aim in these situations is typically to find a way to showcase a particular entity, and a user may be able to benefit from the factors that affect a metric besides performance; for instance, a big institution may choose to use one of the "Power Metrics" that tend to increase as the entity gets bigger, whereas a small institution may choose to use a size-normalized metric

- **Scenario modeling**, such as that which supports the decision of which academic to recruit to an existing research team, or the thinking behind reorganizing a school. The importance of factors besides performance that affect the values of metrics may or may not be important, depending on the particular scenario that is being modeled.

3.2 Factors besides performance that affect the value of a metric

Six types of factors, besides performance, that may affect the value of a metric are discussed in this section:

- Size

- Discipline

- Publication-type

- Database coverage

- Manipulation

- Time

In situations where a metric does not inherently account for a variable that may be important to address particular questions, SciVal supports the user by providing options in terms of functionality and other metrics; these options are outlined in the sections of this Guidebook dedicated to each individual metric.

	Size-normalized?	Field-normalized?	Publication-type normalized?	Resistant to database coverage?	Difficult to manipulate?	Time-independent?
Academic-Corporate Collaboration	½				✓	✓
Academic-Corporate Collaboration Impact	✓					
Awards Volume					✓	✓
Citation Count						
Citations Per Publication	✓					
Cited Publications	½					
Citing-Patents Count					✓	
Collaboration	½				✓	✓
Collaboration Impact	✓					
Field-Weighted Citation Impact	✓	✓	✓	✓	✓	
Field-Weighted Mass Media	✓	✓	✓	✓	✓	✓
Field-Weighted Views Impact	✓	✓	✓	✓	✓	✓
h-indices						
Mass Media					✓	✓
Media Exposure	✓				✓	✓
Number of Citing Countries					✓	
Outputs in Top Citation Percentiles	½	½				
Outputs in Top Views Percentiles	½	½				✓
Patent-Citations Count						
Patent-Citations per Scholarly Output	✓					
Patent-Cited Scholarly Output						
Publications in Top Journal Percentiles	✓	✓			✓	✓
Scholarly Output					✓	✓
Scopus Source Title Count					✓	✓
Subject Area Count					✓	✓
Views Count						✓
Views per Publication	✓					✓

Table 1: Characteristics of the metrics in SciVal. Those metrics that have half a shaded cell in the "Size-normalized" column are size-normalized when the "Percentage" option is selected, but not when the "Total value" option is selected.

ELSEVIER

3.2.1 Size

There are some metrics whose value tends to increase with the size of an entity, such as Scholarly Output that indicates the productivity of an entity, and Citation Count that sums, over all publications, the citations received by an entity. These metrics are referred to within this Guidebook as "Power Metrics", and are summarized in Table 1.

It is often important to account for differences in the size of an entity when evaluating performance; Citations per Publication, for instance, accounts for differences in the size of an entity's Scholarly Output, and is useful to reveal efficiency of citations received per item. When demonstrating excellence, however, "Power Metrics" may be used to the benefit of large entities whose performance will tend to look better; Citation Count will generally be higher for a large collaboration network than for a small research team.

3.2.2 Discipline

Academics working in different disciplines display distinct characteristics in their approach to research and in their communication about research. These behavioral differences are not better or worse than each other, but are merely a fact associated with particular fields of research.

Any Citation Count-type or Citations per Publication-type metric effectively illustrates these differences.

These metrics tend to be significantly higher in Neuroscience than in Engineering, for example, but it is obviously not the case that Neuroscience is generally "better" than Engineering. These types of metric do not take into account different behaviors between fields, and it is not advised to use them to make comparisons between fields: for this purpose, field-normalized metrics, such as Field-Weighted Citation Impact, Publications in Top Journal Percentiles, and the free journal metrics SNIP and SJR[11], are suitable.

It is not only Citation Count-type and Citations per Publication-type metrics that are affected by these behavioral differences; they affect all metrics used throughout research management, including those implemented in SciVal. What is it that causes these differences?

- **Frequency of publication.** Academics in fields such as chemical engineering tend to publish more often than researchers in mathematics.

- **Length of reference lists.** Publications in disciplines such as toxicology tend to have much longer reference lists than those in social sciences.

- **Number of co-authors.** Research in Physics tends to be much more collaborative than research in arts and humanities, resulting in a higher number of co-authors per publication.

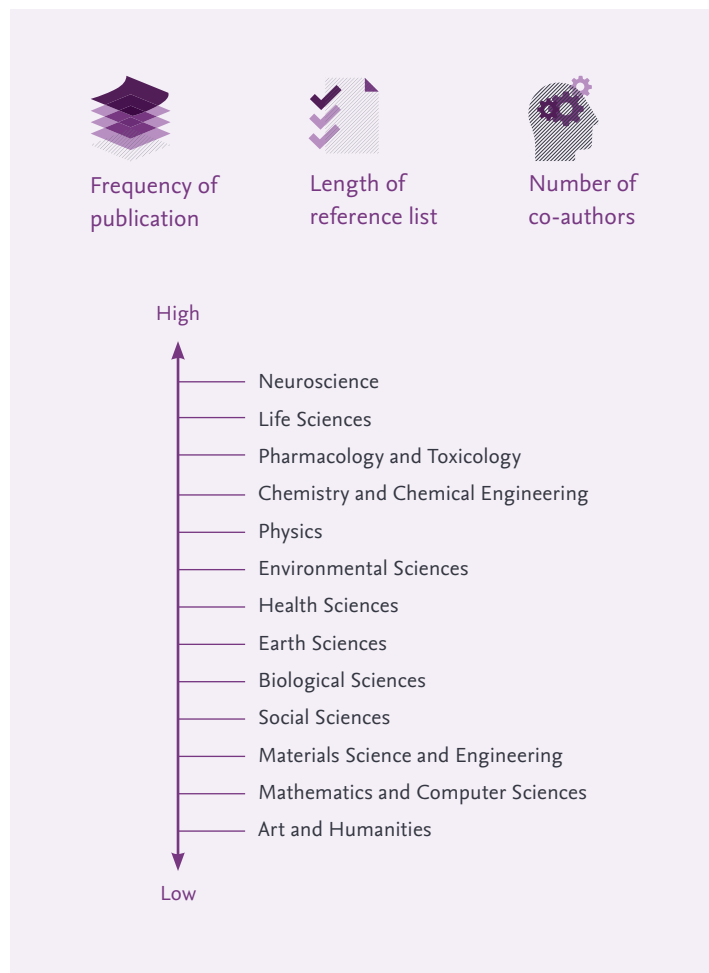The typical distribution of these behaviors amongst all disciplines is illustrated in Figure 1.

11 https://www.scopus.com/sources

Figure 1: The characteristic behavior of academics differs between disciplines

What does this mean for metrics? The majority of metrics values tend to be higher in Neuroscience and Life Sciences than they are in Materials Science and Computer Sciences; this does not necessarily reflect performance, but merely the way that academics in these disciplines conduct research. This may be used to advantage when the aim is to demonstrate excellent performance of an academic working in Neuroscience, for instance, but it is very important to be aware of this effect and to take it into account when performing an evaluation.

If users need to make comparisons between disciplines, they are advised to:

- **Apply the Research Area filter** when comparing entities made up of a mixture of disciplines, such as an Institution or a Country, to focus on one field that is common between all the entities

- **Select field-normalized metrics**, such as Field-Weighted Citation Impact

It is worth noting, when considering the effect of disciplines on metric values, that there are different ways of defining disciplines:

- **Journal classifications** are probably the most common approach to defining disciplines

  - Scopus, for example, assigns the journals that it indexes into Main Categories and Sub-Categories, and the Brazilian CAPES Foundation[12], a government agency that awards scholarship grants to graduate students at universities and research centers in Brazil and abroad, also has a journal classification, as do many other organizations

  - These classifications offer the significant advantage to users of a tool like SciVal of having the same meaning, regardless of the entity being investigated. The meaning of a Chemistry department may differ between research institutions, and that is a concern when benchmarking Chemistry. If a user filters institutions by a particular journal category they can be sure that they are comparing a consistent definition of a discipline. It is for this reason that the Subject Area filter in SciVal offers journal classifications

  - The drawback of these classifications is that they tend to be very large so that it is difficult to keep them up to date with current developments in research, and that it is assumed that a journal's discipline(s) applies equally to each individual publication that it contains

- **Publication-level classifications** are gaining increased attention, as the pace of change in academia increases and with technological advances that can handle massive data sets:

  - These classifications do not assume that a publication fits into the same discipline(s) as the journal that it appears in. They are often formed by allowing publications to group themselves through the references they contain, and/or by the publications that refer to them

  - Publication-level classifications are very useful to keep pace with the changing ways in which academics view the world of research, and can be useful to reveal nascent, emerging fields. The Topics of Prominence in SciVal are examples of research activity that represent the way that recent publications cite previous work. It is also for this reason that publication-level classifications are used to determine disciplines in the calculation of Field-Weighted Citation Impact in SciVal

  - The drawbacks of these classifications are:

    - There tend to be publications which cannot be easily assigned, for instance if they have a short or absent reference list and/or have not yet received citations

    - It is computationally very demanding to implement an ever changing classification

    - Frequent changes in classification due to changes in the pattern of citations received reduce the transparency of the data underlying a metric calculation for users

- **User-defined disciplines**, for instance by searching for key phrases within the titles and abstracts of a publications database:

  - The advantage of user-defined disciplines is that they ensure that every user can define the world of research as it makes sense to them and their questions. Academics do not conduct their research in a way that conveniently fits the journal classifications and previous citing behavior of their colleagues, and even publication-level classifications are not able to detect the very first stages of Subject Areas, when academics are just starting to focus on a new field

  - The drawback of user-defined disciplines is that they are likely unique, and pose challenges to others who would like to use or validate these customized fields

3.2.3 Publication-type

Different types of publications tend to be cited with different rates. The most well-known example is that reviews tend to attract more citations than original research articles, but there are variations between other publication-types as well. These are illustrated in Figure 2.

Publication-type differences may not be important where the performance of total output is of interest, or when a journal editor is showcasing their productivity and wishes to retain the positive effect of editorials in the count. There are some situations, however, when there may be a strong preference to focus on a particular type, such as on conference proceedings in engineering and computer science, or original research articles when comparing a journal editor who has written many editorials with an academic who is not a journal editor.

If this effect is important for the question being addressed by the user, they are advised to:

- Apply the Publication-type filter

- Use **Field-Weighted Citation Impact** which is normalized for Publication-type

12 http://www.iie.org/en/Programs/CAPES

Citations per Publication
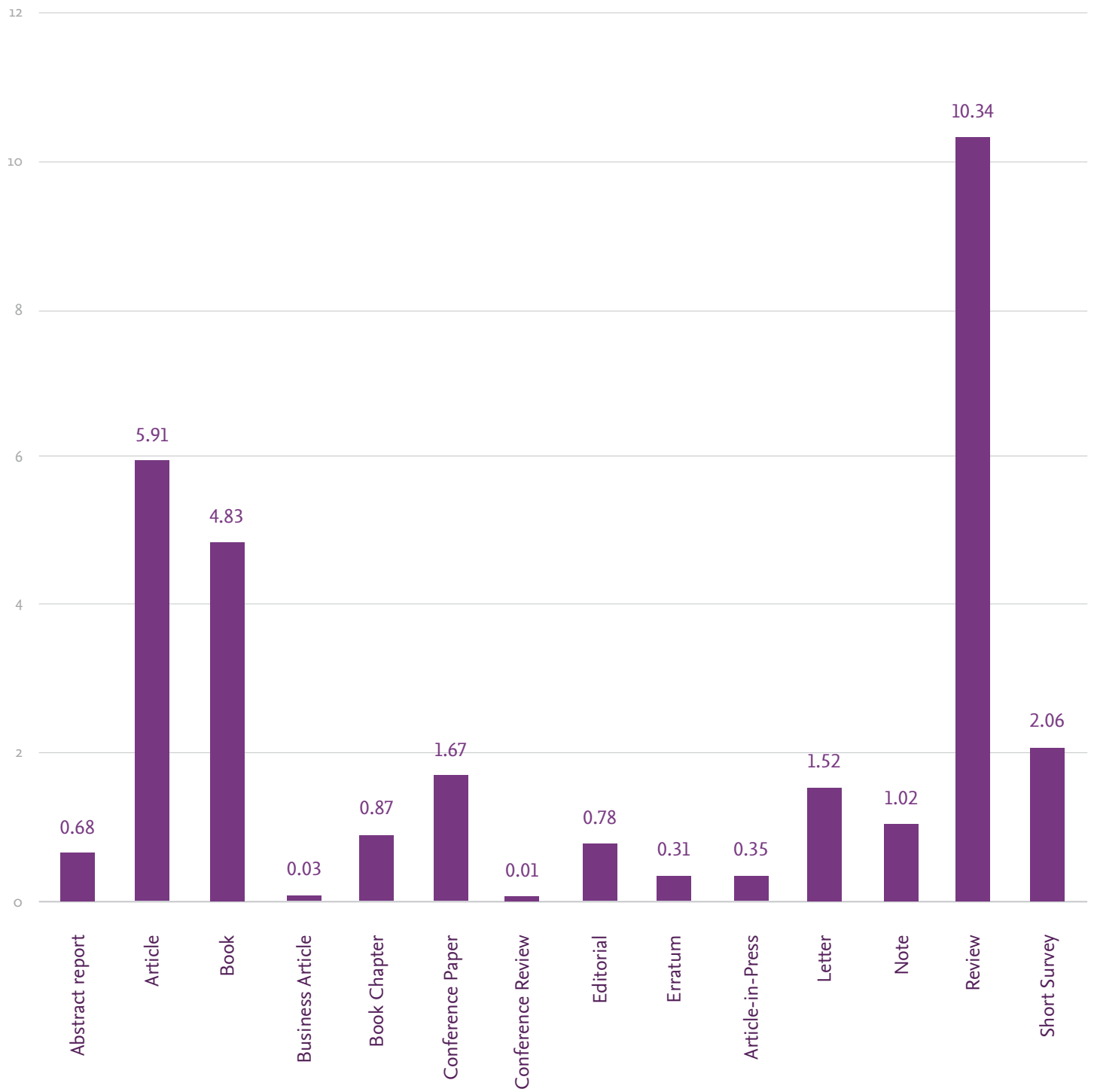
Figure 2: Citation rates for different publication-types as classified in Scopus. This chart displays citations received up to May 2018 per item published during the period 2013-2017.

3.2.4 Database coverage

Databases have particular guidelines in determining which content to include. Scopus has a comprehensive policy to select the content which meets its aims, and the policy[13] is to be selective and not to include every single publication globally. This means that there may be some items that have been published by a particular entity that are not indexed in Scopus, and so cannot be part of the metrics calculations in SciVal.

There are 2 aspects to considerations of database coverage:

- **Geographical coverage.** Scopus indexes content from more than 5,000 publishers from all over the world. The geographical distribution of titles indexed in Scopus is representative of the global concentrations of publishers, with the focus of activity in the United States and the United Kingdom, as shown in Figure 3. This geographical coverage should support a thorough analysis of topics of global interest; however, for research areas of primarily local interest, such as national literature, history or culture, Scopus may not provide sufficiently complete data

- **Disciplinary coverage.** The ongoing expansion of the titles indexed by Scopus means that this coverage will continue to change. The disciplinary coverage of Scopus can be estimated by looking at the items that have been cited by recently published work; the extent to which these citations can be linked to items indexed within the Scopus database represents the coverage, and those citations which refer to items not indexed by Scopus are assumed to represent lack of coverage. This methodology represents a minimum coverage level, and the actual coverage is probably a few percentages higher than shown in Figure 4.

13 http://www.elsevier.com/online-tools/scopus/content-overview

Figure 3:
Geographical distribution of titles indexed in Scopus, to 30th April 2018. The country assigned to a title is the country of the publisher imprint.
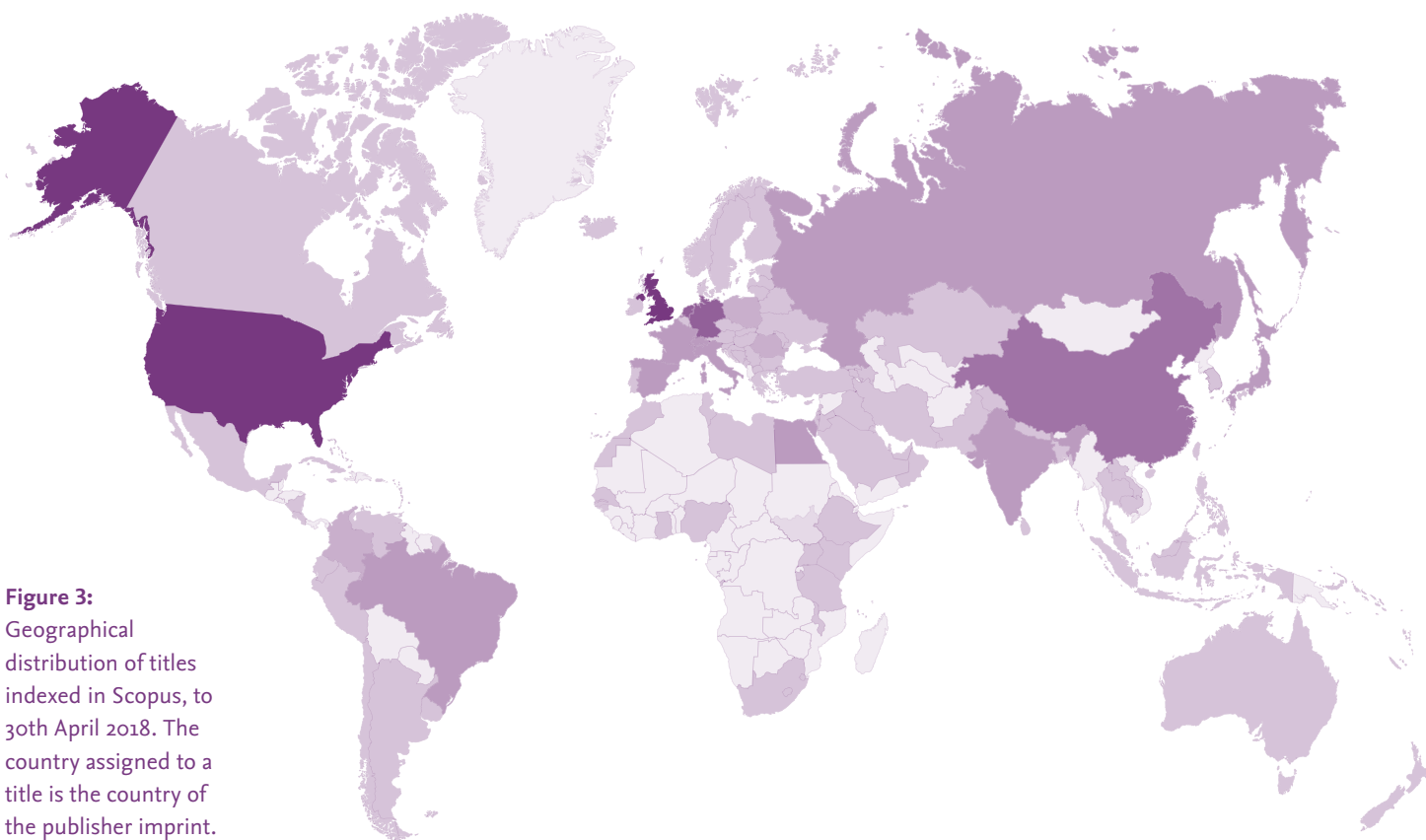
The sliding scale indicates the density of indexed titles.

6,000+   0

ELSEVIER

Citations linking to items outside of Scopus
or to pre-1996 items in Scopus

Citations linking to post-1995 items in Scopus
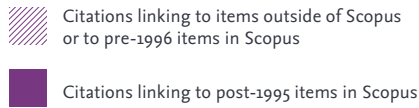
Figure 4: Estimation of Scopus' disciplinary coverage. This estimation is made based on the extent to which citations from publications during the period 2013-2017 can be linked to items indexed from 1996 onwards within the Scopus database; this is an under-estimation of the true coverage because citations to indexed items published before 1996 are not captured here. This analysis is based on data up to May 2018.

It is most probably acceptable to have a few missing publications from an entity's output when benchmarking large data sets against their peers, because gaps in database coverage likely affect all entities similarly and will not invalidate the comparison. Care is advised, however, when comparing small entities, from which a single missing publication may have a significant negative impact; for example, an academic's performance may suffer due to gaps in database coverage of their portfolio, as well as gaps in publications citing those items of the portfolio that are indexed. The transition point between a "small" and a "large" entity is a matter of judgment, and will differ depending on the discipline. The only way to account for this is to be vigilant and to apply common sense when using SciVal to support decisions.

The question of the effect of database coverage is most often raised in relation to Arts and Humanities, and Social Sciences: are the metrics in SciVal useful in these fields?

In some situations, they are:

- The type of decisions supported by SciVal should always be based on a combination of inputs from peer review and expert opinion, as well as from quantitative evidence such as metrics. Publication and citation data can therefore form part of the evidence-base, with funding, innovation, and societal impact, for instance, also being very important and partially addressed in SciVal. This applies across all disciplines, including Arts and Humanities and Social Sciences

- Another concern in these fields is around their coverage in Scopus and other commercial databases which tends to be lower than other disciplines. This is in part a natural consequence of the publishing behavior of academics in these fields, who favor the publication of stand-alone books which are difficult to capture relative to serials, although Scopus is now focusing on increasing its coverage of these types of output. Nevertheless, valuable information about performance in these fields can be gained from SciVal if the guidelines about the size of the entity noted above are appropriately considered

3.2.5 Manipulation

Some situations lend themselves relatively easily to manipulation for the artificial inflation of metrics values. One example is the combination of research units to artificially increase size for reporting purposes, which tends to improve apparent performance when using "Power Metrics".

Another example is that of self-citations. There is nothing inherently wrong about self-citing: it is normal academic behavior to build on work previously published, and it is an author's responsibility to refer a reader to older material that will support their understanding. This practice is, however, open to abuse by unscrupulous academics who could choose to cite irrelevant past papers to boost their citation counts, and to journal editors who occasionally coerce authors of submitted papers to add too many additional citations to publications within their journal to their reference lists.

Citation Impact metrics are the most susceptible to manipulation of the metrics in SciVal. Self-citation abuse is very rare, but users have the option to exclude self-citations from several of these metrics if they

wish, as illustrated in Example 1. Other metrics are much more difficult to manipulate, such as those that provide information on aspects of productivity and collaboration.

3.2.6 Time

The passage of time is critical to enable useful information to be derived from some metrics. The most obvious of these are Citation Impact metrics, since time is needed for published work to receive citations. The *h*-indices are another example of metrics that are not very useful when applied to the output of an early-career researcher, because the need for the passage of time to accumulate citations is coupled with that for a reasonably-sized body of work to provide reliable information.

Some "time-independent" metrics can be used immediately upon entry of a publication into the database. Examples are metrics that look at collaboration by using affiliation information. There is even a Citation Impact metric that is independent of time; Publications in Top Journal Percentiles uses citations received by the journals in which publications appear.

4.0 SciVal and research metrics

4.1 Groups of metrics in SciVal

SciVal offers a broad range of metrics to:

- Accommodate the preferences of users in approaching questions from multiple angles

- Enable users to "triangulate" their evidence. When seeking quantitative input into a question through research metrics, you should use at least 2 different metrics and where pieces of intelligence gained from multiple metrics reinforce each other, this can provide the user with a higher degree of confidence that their conclusions are valid

It is possible to classify metrics in a variety of ways. The majority of SciVal's metrics can be classified within 6 groups, and a metric may be part of more than 1 group, as illustrated in Table 2:

- **Productivity metrics** give information on the volume of output of an entity

- **Citation Impact** metrics indicate the influence of an entity's output, as indicated by various types of citation counts

- **Collaboration metrics** provide information on the research partnerships of an entity

- **Disciplinarily metrics** give information on the spread of topics within an entity's publications

- **Snowball Metrics**[14] are defined and endorsed by research-intensive universities as providing important insight into institutional strategies:

  – The agreed and tested definitions are shared free-of-charge with the sector in the Snowball Metrics Recipe Book[15], with the ambition that Snowball Metrics become global standards for the higher education sector

  – These "recipes" can be used by anyone for their own purposes. Elsevier supports Snowball Metrics as recognized industry standards, and is implementing these metrics in relevant systems and tools, including SciVal

  – Snowball Metrics are indicated in the product interfaces by the following snowflake symbol

- **"Power Metrics"** whose value tends to increase as the size of an entity increases. For example, a larger institution will tend to publish more output than a smaller institution, just because of its greater size

14 www.snowballmetrics.com
15 www.snowballmetrics.com/metrics

View SciVal metrics overview table on the next page

→

	Citation Impact	Collaboration	Snowball Metric	"Power Metric"	Views Metric	Funding Metric	Patent Metric	Media Metric
Academic-Corporate Collaboration		■	■					
Academic-Corporate Collaboration Impact	■	■	■					
Awards Volume			■	■		■		
Citation Count	■		■	■				
Citations per Publication	■		■					
Cited Publications	■			▨				
Citing-Patents Count	■						■	
Collaboration		■	■					
Collaboration Impact	■	■						
Field-Weighted Citation Impact	■		■					
Field-Weighted Mass Media								■
Field-Weighted Views Impact					■			
h-indices	■		■	■				
Mass Media				■				■
Media Exposure				■				■
Number of Citing Countries	■			■				
Outputs in Top Citation Percentiles	■		■	▨				
Outputs in Top Views Percentiles				▨	■			
Patent-Citations Count	■			■			■	
Patent-Citations per Scholarly Output	■						■	
Patent-Cited Scholarly Output	■			■			■	
Publications in Top Journal Percentiles	■		■	▨				
Scholarly Output			■	■				
Scopus Source Title Count				■				
Subject Area Count				■				
Views Count				■	■			
Views per Publication					■			

Table 2: Overview of metrics in SciVal. Those metrics that have half a shaded cell in the "Power metric" column are "Power metrics" when the "Absolute number" option is selected, but not when the "Percentage" option is selected.

ELSEVIER

It is also possible to classify metrics based on what they are providing insight into. Within SciVal, the metrics have therefore been grouped into 7 categories to help users identify suitable metrics more easily as shown in Table 3.

	Awarded Grants	Collaboration	Published	Viewed	Cited	Economic Impact	Societal Impact
Academic-Corporate Collaboration		■				■	
Academic-Corporate Collaboration Impact		■			■	■	
Awards Volume	■						
Citation Count					■		
Citations per Publication					■		
Cited Publications					■		
Citing-Patents Count					■	■	
Collaboration		■					
Collaboration Impact		■			■		
Field-Weighted Citation Impact					■		
Field-Weighted Mass Media							■
Field-Weighted Views Impact				■			
h-indices			■		■		
Mass Media							■
Media Exposure							■
Number of Citing Countries					■		
Outputs in Top Citation Percentiles					■		
Outputs in Top Views Percentiles				■			
Patent-Citations Count					■	■	
Patent-Citations per Scholarly Output					■	■	
Patent-Cited Scholarly Output					■	■	
Publications in Top Journal Percentiles					■		
Scholarly Output			■				
Scopus Source Title Count			■				
Subject Area Count			■				
Views Count				■			
Views per Publication				■			

Table 3: Grouping of metrics in SciVal based on what they provide insight into.

ELSEVIER

4.2 The calculation and display of metrics in SciVal

4.2.1 Publications included in the calculation of a metric

The ideal situation would be that every publication in a data set is associated with the information needed so that it can be included in the calculation of every metric. In practice this is not the case; authors do not always include complete affiliation information, and publications are not always part of items indexed in Scopus that have journal metrics values, for example. **Publications that lack the necessary information are excluded from the metric calculation.**

4.2.2 Deduplication

SciVal offers the user the opportunity to investigate aggregate entities formed by the combination of smaller entities; for example, US states are Groups of Institutions, and geographical regions are Groups of Countries.

The same publication may be part of multiple smaller component entities, and can be added multiple times to the aggregate entity. Say, for example, that the Researcher R1 has co-authored a publication P1 with Researcher R2; in that case, P1 is a part of both R1 and R2.

SciVal deduplicates all the publications within an aggregate entity, so that a publication is only counted once even if it is co-authored by several of the component entities. **Entities are groups of unique publications**, and users can have confidence when creating aggregate entities that SciVal deduplicates the data set. In the above example, P1 is counted once only in a Group of Researchers that is composed of R1 and R2.

4.2.3 Zero and null values

SciVal distinguishes between zero and null values (absence of a value), which have distinct meanings for the interpretation of metrics.

Consider the metric Scholarly Output, which calculates the number of publications of an entity. If an entity has not published any outputs in a particular time frame, this is displayed as a zero value and the lack of publication is important to understanding an entity's activity. Null values are never displayed for Scholarly Output.

This is not the case for a metric like Citation Count, which counts the number of citations that an entity's publications have received. It is not meaningful to display a value of zero citations for a period during which an entity has not published anything, because the entity simply cannot receive citations when there is an absence of publications. This is different from the situation when an entity has published items that have not received any citations.

SciVal therefore distinguishes between the following 2 cases:

- If an entity has published output in a particular period, and that output has not received any citations, then a value of zero is displayed for Citation Count

- If an entity has not published any output in a particular time frame, then the entity cannot receive any citations during that period. In this case, a null value is plotted in SciVal; users may notice a break in a line when viewing metrics over time in the Benchmarking module, for example

This same reasoning is applied to other metrics. It is not possible to be internationally collaborative, for example, if nothing has been published.

Outputs in Top Citation Percentiles can only be calculated for the current year from the first data snapshot on or after 1 July. It will be displayed as a null value until this date is reached. This metric depends on being able to divide the publications into 100 percentiles, and this level of division is not possible earlier in the publication year when items just published have received very few citations.

4.2.4 The display of ">current year"

A high proportion of the items indexed in Scopus are journals. Publishers sometimes release journal issues before their official cover date; for instance, at the end of 2018, users can find items in the Scopus database that belong to the 2019 publication year, and occasionally even to later years.

A great deal of importance is placed in SciVal on a high currency of data, so that all information present in the database is available to users. All publications that belong to items with a cover date later than the current year are collected together in the "Publication Year" labelled ">2018" in 2018, ">2019" in 2019, and so on.

4.2.5 Citation Counts

Citations Counts in SciVal represent the total number of citations received since an item was published, up to the date of the last data cut. Citation Impact metrics are often displayed in a chart or table with years in SciVal to indicate trends; these years are always the years in which items were published, and never refer to the years in which citations were received.

Older publications tend to have more citations than newer publications, simply because they have had longer to receive them from subsequent work. This means that for metrics like Citation Count and Citations per Publication, where this time factor is not accounted for, the user will notice a "dip" in the timeline in recent years. Comparisons between similar entities will still be useful and valid despite this dip, since the time factor will affect all similar entities equally, but if the user prefers to avoid this display they can select a metric like Field-Weighted Citation Impact which inherently accounts for this.

4.2.6 Calculation options

SciVal users have different needs in how they use the metrics. Some of these needs are met by how the results are visualized, but some affect the method of calculation. For instance:

- An academic who is a journal editor may well publish regular editorials, which count towards their productivity and may also increase their citation impact. In understanding the citation impact of this

academic, it is very important to look at the entire body of their output, including the editorials, since being a journal editor and writing editorials is a significant part of their scholarly contribution. However, when comparing this academic to others who are not journal editors, it may sometimes be desirable to exclude the effect of the editorials to ensure a fair comparison

- Self-citations are those by which an entity refers to its previous work in new publications. Self-citing is normal and expected academic behavior, and it is an author's responsibility to make sure their readers are aware of related, relevant work. However, some users prefer to exclude self-citations from Citation Impact metrics, whereas to others this is not critical to the analysis being done

There is no universally right or wrong approach to these calculation preferences, although in some situations an option may become very important. These calculation preferences are available in SciVal as options to the user, some of which are described below. Each metric has its own set of calculation options, which are summarized in Table 4.

4.2.6.1 Subject Area filter

The Subject Area filter limits the metric calculations to the publications of an entity that fall within a particular discipline. Where journal classifications are used, as in the accompanying screenshot, both the Main Categories and Sub-Categories are available to use as Subject Area filters. User-defined Research Areas can also be used in the Collaboration and Trends modules, for example the areas of "Graphene" and "Malaria".

The Subject Area filter applies to the entity's publications only. It does not have any effect on the citation counts used to calculate the Citation Impact metrics; the citations are counted regardless of the Subject Area of the citing paper.

Some metrics are not field-normalized. This means that differences in the behavior of academics in distinct disciplines that can affect metrics values are not accounted for in their calculation, and it may be difficult for a user to distinguish between differences in disciplinary behavior and true differences in research activity. These non-field-normalized metrics are very useful when comparing entities that fall into the same discipline, but it is not advisable to use non-field-normalized metrics to compare entities that fall into different fields.

Users can address this by:

- Using the Subject Area filter when using a non-field-normalized metric to compare entities made up of a mixture of disciplines, such as an Institution or a Country. This has the effect of enabling benchmarking of comparable disciplinary slices of these entities

- Using field-normalized metrics such as Field-Weighted Citation Impact which inherently take disciplinary differences into account

The default setting in SciVal is that no Subject Area filter is applied.

4.2.6.2 Publication-type filter

The Publication-type filter limits the metric calculations to articles, reviews, and/or books for example. This filter can be applied when the user judges that this variable is important for their analysis, such as:

- Distinguishing between original research contributions that are often published as articles, and expert opinion, typically communicated in reviews

- In disciplines such as Engineering and Computer Science, where it is sometimes important to focus on conference proceedings

- When comparing an academic who is a journal editor and has the opportunity to publish editorials, with an academic who is not a journal editor
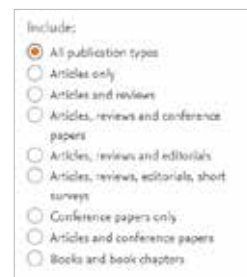
The Publication-type filter applies to the entity's publications only. It does not have any effect on the citation counts used to calculate the Citation Impact metrics; the citations are counted regardless of the publication-type of the citing paper. The only exception is Outputs in Top Citation Percentiles, where this filter also applies to the data universe used to generate the citation thresholds.

The default setting in SciVal is that no Publication-type filter is applied.

4.2.6.3 Self-citation exclusion

Self-citations are those by which an entity refers to its previous work in new publications. Self-citations are typically thought of in the context of an individual academic, but journal self-citations, institutional self-citations, and country-self-citations are examples of the same activity at different levels.

There is nothing inherently wrong with the act of self-citing. It is normal and expected academic behavior to build on work previously conducted and published, and to refer in new publications to older material to alert the reader to important and relevant information that will aid their understanding. Indeed, the most relevant work to be aware of may well be research that has previously been published by the same entity, and it is an author's responsibility to make their readers aware of it.

The act of referring to previous work is, however, open to abuse, and there are rare cases of unethical self-citation practices which have led to self-citation acquiring a somewhat bad reputation. SciVal users have the option to exclude self-citations from the calculation of many of the Citation Impact metrics, to make their own judgment about whether self-citation rates are within normal boundaries. SciVal distinguishes between self-citations at the author, institutional, and country levels, and applies the appropriate exclusion depending on the type of entity being viewed without the user needing to specify this.

The default setting in SciVal is that all citations are included.

See Examples 1a and 1b for illustrations of how self-citation exclusion is performed in SciVal.

4.2.6.4 Total value and percentage options

Some metrics in SciVal offer the option of being viewed as "Total value", or as "Percentage":

- Users are advised to select the "Percentage" option to size-normalize, when they are comparing entities of different sizes.

- It is useful to select the "Total value" option when it is important to communicate the magnitude of the number of publications involved.

The default setting in SciVal is "Percentage".

	Subject Area filter	Publication-type filter	Self-citation exclusion	Total value or Percentage	Field weighted	Collaboration	Percentile level	Journal metric	h-indices	Media type	Funding organization	Patent office
SciVal default	No filter applied	No filter applied	No exclusion	Percentage	No	International	10%	CiteScore	h-index	print	all	all
Academic-Corporate Collaboration	✓	✓		✓								
Academic-Corporate Collaboration Impact	✓	✓										
Awards Volume	✓										✓	
Citation Count	✓	✓	✓									
Citations Per Publication	✓	✓	✓									
Cited Publications	✓	✓	✓	✓								
Citing-Patents Count	✓											✓
Collaboration	✓	✓		✓	✓	✓						
Collaboration Impact	✓	✓				✓						
Field-Weighted Citation Impact	✓	✓	✓									
Field-Weighted Mass Media	✓									✓		
Field-Weighted Views Impact	✓	✓										
h-indices	✓	✓	✓						✓			
Mass Media	✓									✓		
Media Exposure	✓									✓		
Number of Citing Countries	✓	✓	✓									
Outputs in Top Citation Percentiles	✓	✓	✓	✓								
Outputs in Top Views Percentiles	✓	✓		✓								
Patent-Citations Count	✓	✓			✓		✓					✓
Patent-Citations per Scholarly Output	✓	✓			✓		✓					✓
Patent-Cited Scholarly Output	✓											✓
Publications in Top Journal Percentiles	✓	✓		✓								
Scholarly Output	✓	✓						✓				
Scopus Source Title Count	✓	✓										
Subject Area Count	✓	✓										
Views Count	✓	✓										
Views per Publication	✓	✓										

Table 4: Calculation options for the metrics in SciVal

Example 1a: Self-Citation Exclusion

Scenario: The user is looking at a Country, Institution or Researcher entity that consists of 1 publication, PP. Say that this entity has received 6 citations from P1, P2, P4, P5, P6 and P7.

Publication	Author	Institution	Country
PP	AA	II	CC

PP is cited by

Publication	Author	Institution	Country
P1	A1	I1	C1
P2	AA	I2	C2
P4	AA	I3	CC
P5	AA	II	CC
P6	A2	II	CC
P7	A2	I3	CC

Question: What happens when the user chooses not to include self-citations when calculating a metric that offers this as an option?

Answer:

If the entity you are viewing is Country CC, then the citations from P4, P5, P6 and P7 are self-citations because the affiliation country CC is the same as the entity's Country CC.

These 4 citations would no longer be included in the metric calculation.

Answer:

If the entity you are viewing is Institution II, then the citations from P5 and P6 are self-citations because the affiliation institution II is the same as the entity's Institution II.

These 2 citations would no longer be included in the metric calculation.

Answer:

If the entity you are viewing is a Researcher that includes author AA, then the citations from P2, P4 and P5 are self-citations because the author AA is the same as that of the entity's author AA.

These 3 citations would no longer be included in the metric calculation.

The citation from P1 is never classified as a self-citation, regardless of the type of entity.

Example 1b: Self-Citation Exclusion

Scenario: The user is looking at a Country, Institution or Researcher that consists of 3 publications, P4, P5 and P6:

- The publication P4 has 2 authors: A2 and A5. They are both based at institution I2, in country C2.

- The publication P5 has 2 authors: A1 and A3. Author A1 has joint affiliations: I1 and I3 in countries C1 and C2 respectively. A3 has a single affiliation: I2 in country C2.

- The publication P6 has the same 2 authors as P5: A1 and A3. Author A1 has again listed joint affiliations: A1 has remained at I3 in country C2, but the second affiliation is I5 in C4. A3 still has a single affiliation but it is different from that on P5: the affiliation on this publication P6 is I5 in country C4.

Publication	Author	Institution	Country
P4	A2	I2	C2
	A5	I2	C2
P5	A1	I3	C2
		I1	C1
	A3	I2	C2
P6	A3	I5	C4
	A1	I5	C4
		I3	C2

P4, P5 and P6
are cited by:

Publication	Author	Institution	Country
P1	A1	I6	C4
	A3	I6	C4
P2	A3	I2	C2
		I3	C2
	A4	I3	C2
P3	A4	I4	C3
	A5	I2	C2
		I6	C4
	A6	I6	C4
	A2	I4	C3

Question: What happens when the user chooses not to include self-citations when calculating a metric that offers this as an option?

Answer:
If the entity you are viewing is, or includes, Countries C1, C2 or C4, then:

- A citation from P1 to P6 is a self-citation because they share the affiliation country C4

- A citation from P2 to all of P4, P5 and P6 is a self-citation: P2 shares affiliation country C2 with each of P4, P5 and P6

- A citation from P3 to all of P4, P5 and P6 is a self-citation: P3 shares affiliation country C2 with each of P4, P5 and

P6, and in addition shares affiliation country C4 with P6

- A citation from P1 to P4 and P5 is not a self-citation because P1 does not share a country with either P4 or P5

Answer:
If the entity you are viewing is, or includes, Institutions I1, I2, I3 or I5, then:

- A citation from P2 to all of P4, P5 and P6 is a self-citation: P2 shares institution I2 with both P4 and P5, and institution I3 with P5 and P6

- A citation from P3 to P4 and P5 is a self-citation: P3 shares institution I2 with both P4 and P5

- A citation from P1 to any of the entity's publications is not a self-citation because it does not share an institution with any of P4, P5 or P6

- A citation from P3 to P6 is not a self-citation because these publications do not have any institution in common

Answer:
If the entity you are viewing is, or includes, authors A1, A2, A3 or A5, then:

- A citation from P1 to both P5 and P6 is a self-citation: P1 shares authors A1 and A3 with both of P5 and P6

- A citation from P2 to both P5 and P6 is a self-citation: P2 shares author A3 with both P5 and P6

- A citation from P3 to P4 is a self-citation: P3 shares

authors A2 and A5 with P4

- A citation from P1 or P2 to P4 is not a self-citation because these publications do not have any author in common

- A citation from P3 to P5 or P6 is not a self-citation because these publications do not have any author in common
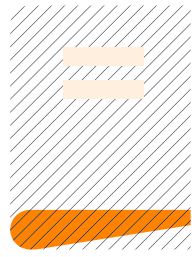
5.0 Research metrics in SciVal: Methods and use

ELSEVIER

5.1 Awarded Grants

5.1.1 Metric: Awards Volume

Awards Volume in SciVal refers to both the count and the value of grant awards. Awards Volume considers aggregated values of awards over the award lifetime. In other words, it considers the total value awarded at the time of award and not the value (to be) spent in any particular time period.

Awards Volume is displayed either as count of awards or the US Dollar value of the awards. The US Dollar values are calculated using the historical exchange rates provided by the US Federal Reserve. Awarded Grants cover years starting from 2009 to present.

Subject areas are assigned to the grants using Fingerprint® technology where an index of weighted concepts are generated for the subject areas and from the abstract of the awarded grant using the Elsevier Fingerprint Engine®. The weighted concepts for the subject areas and grants are then matched with each match resulting in that award being assigned to that subject area. Each award can be assigned to multiple subject areas.

Grants are assigned to institutions based upon the Scopus affiliation of the principal investigator (PI) at the time of the grant being awarded. Institutions in SciVal are based on Scopus affiliations. Countries are assigned based upon the country issuing the grant. Institutions in SciVal can be made up of multiple Scopus affiliations, but will only be assigned to a single country.

When calculating Awards Volume, SciVal considers both the institution and the country of the awarded grant. Each Scopus affiliation has a country and in cases of institutions with overseas branches, an institution in SciVal can consist of affiliations assigned to more than one country. When Award Volume is calculated on country level the Scopus affiliation is used to decide the country assigned. That way all the grants awarded to PI's working in a particular country are captured and assigned correctly. When Awards Volume is calculated on an institutional level all grants assigned to that SciVal institution are used.

SciVal gathers award data from the following funding bodies:

- UK
    - Wellcome Trust (WT)
    - Arts and Humanities Research Council (AHRC)
    - Biotechnology and Biological Sciences Research Council (BBSCR)
    - Economic and Social Research Council (ESRC)
    - Engineering and Physical Sciences Research Council (EPSRC)
    - Medical Research Council (MRC)
    - Natural Environment Research Council (NERC)
    - Science and Technology Facilities Council (STFC)

- US
    - National Institutes of Health (NIH)
    - National Science Foundation (NSF)
- Australia
    - Australian Research Council (ARC)
    - National Health and Medical Research Council (NHMRC)

Awards Volume is a:

- Snowball Metric
- "Power Metric"

This metric is useful to:

- Understand how much funding the principal investigators of an organization receive from various funding bodies:
    - Since funding bodies usually do not share what publications were produced by the grants they awarded, SciVal does not provide this information at the publication level.
- Benchmark similar institutions (size and country) to understand how an institution compares in awards funding to their peers
- Showcase funding received

This metric should be used with care when:

- Benchmarking between organizations with different profiles:
    - It is better to compare entities from the same country as funding bodies are often country specific.
    - The size of the institution affects how many awards they receive. Smaller institutions should not be compared to larger ones.
- Benchmarking entities with different disciplinary profiles:
    - Different disciplines have very different funding patterns. When comparing the funding of different entities, it is best to compare within a specific discipline.

Useful partner metrics are:

- The number of awards and the US Dollar amount of awards when evaluating or benchmarking funding

5.2 Collaboration

5.2.1 Metric: Collaboration ❄

Collaboration in SciVal indicates the extent to which an entity's publications have international, national, or institutional co-authorship, and single authorship.

Each publication is assigned to 1 of 4 mutually exclusive collaboration types, based on its affiliation information: international, national, institutional, or single authorship. A single publication may of course display each of international, national and institutional collaboration in its affiliation information, but a single collaboration type is assigned

to ensure that the sum of an entity's publications across the 4 categories adds up to 100% of the publications with the necessary affiliation information.

The assignment of geographical collaboration type is performed using the following cascading decision tree:
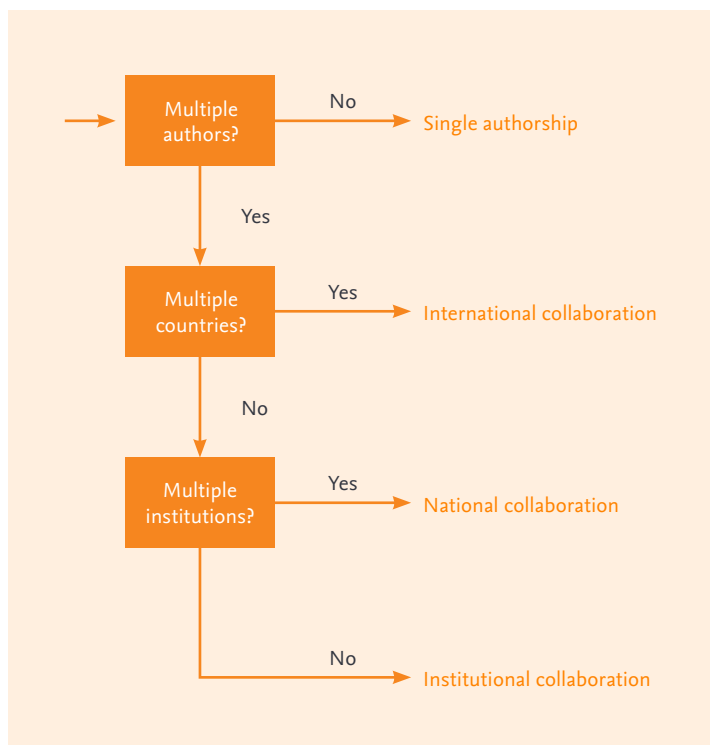
In some cases, a researcher can have two affiliations against their name on a publication. This is not an error, but merely reflects the institutions to which the researcher was affiliated when writing the paper. How does this affect collaboration in SciVal? For example, a researcher publishes an output with both an academic and a corporate institution listed against their name. If they are the sole researcher on the output, then no collaboration is recorded (single authors cannot collaborate with themselves). If they are co-authored with someone affiliated to a different academic institution to the researcher, then academic-to-academic collaboration and academic-to-corporate collaboration is recorded.

When field-weighting Collaboration, a score is calculated using the same methodology as for the calculation of the Field-Weighted Citation Impact. The document level international/national collaboration ratio is computed based on the expected international/national Collaboration for that document type, publication year grouping and subject area assignment. The option to field-weight is only available for Collaboration on the international and national level.

- Field-weighted Collaboration of 1.00 indicates that the entity's Collaboration has been exactly as would be expected based on the global average for similar publications; the Field-Weighted Collaboration of "World", or the entire Scopus database, is 1.00.

- Field-weighted Collaboration of more than 1.00 indicates that the entity's Collaboration has been more than would be expected based on the global average for similar publications; for example, 2.11 means 111% more than the world average.

- Field-weighted Collaboration of less than 1.00 indicates that the entity's Collaboration has been less than would be expected based on the global average for similar publications; for example, 0.87 means 13% less than the world average.

Collaboration is a:

- Collaboration metric

- Snowball Metric

- "Power Metric" when the "Total value" option is selected, but not when the "Percentage" option is selected

This metric is useful to:

- Explore the extent of international and other types of collaboration within a data set

- Benchmark the collaboration of entities of different sizes, but in similar disciplines:
    - It is advised to select the "Percentage" option when comparing entities of different sizes, to normalize for this variable.

- Showcase extensive international collaboration that may underpin a set of scholarly output

- Investigate collaborative activity very early in a new strategy, or for an early-career researcher, since the affiliation data underlying this metric do not require time to accumulate to reliable levels in the same way as citation data do

- Look at publishing activity in a way that is difficult to manipulate

This metric should be used with care when:

- Benchmarking the collaboration of entities in different disciplines:
    - The typical collaborative behavior of academics may differ between disciplines, such as mathematics and medicine, or humanities and molecular biology.
    - It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences.
    - When comparing entities made up of a mixture of disciplines, such as an Institution, or a Group of Institutions such as a US state, it is advised to apply the Research Area filter to focus on one field that is common between all the entities. It is also advised to use field weighting to help offset this affect.

16 https://www.elsevier.com/en-gb/solutions/elsevier-fingerprint-engine
17 https://www.researchtrends.com/issue-39-december-2014/
     field-weighted-internationalization-score/

- Entities are small and there may be gaps in their output within the Scopus coverage:

  – A single missing publication from a small data set may have a significant negative impact on apparent partnerships, whereas the effect of 1 or a few missing publication(s) from a large data set may be acceptable.

  – The only way to account for this is to be vigilant, particularly when looking at small data sets such as the Scholarly Output of an early-career researcher, or to limit the use of Collaboration to comparing larger data sets in which the potential gaps in the database coverage likely have a similar effect on all entities being viewed and do not invalidate the comparison.

- Understanding activity in a discipline with an obvious national focus, such as Finnish literature or cultural studies, when institutional and national collaboration may be of more use than international collaboration.

Useful partner metrics are:

- **Number of Citing Countries**, which indicates the diversity of countries in which Researchers have built on an entity's publications. It is likely to be higher if an entity's international collaboration activity involves several countries, rather than just 1 or 2, even though both situations can give 100% international collaboration.

- **Collaboration Impact**, which calculates the average Citations per Publication for publications with different types of geographical collaboration, and indicates how beneficial these collaborations are with respect to citation impact

- **The set of all other "Power Metrics"** whose value tends to increase as the entity becomes bigger: Scholarly Output, Subject Area Count, Scopus Source Title Count, Citation Count, Cited Publications ("Total value"), Number of Citing Countries, Academic-Corporate Collaboration ("Total value"), Outputs in Top Citation Percentiles ("Total value"), Publications in Top Journal Percentiles ("Total value"), and *h*-indices

- **The set of all other "time-independent metrics"** which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Subject Area Count, Scopus Source Title Count, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

Collaboration, Collaboration Impact, Academic-Corporate Collaboration, and Academic-Corporation Collaboration Impact

5.2.2 Metric: Collaboration Impact ❄

Collaboration Impact in SciVal indicates the citation impact of an entity's publications with particular types of geographical collaboration: how many citations do this entity's internationally, nationally, or institutionally co-authored publications receive, as well as those with a single author?

Publications are assigned to 1 of 4 mutually exclusive geographical collaboration types, as explained for Collaboration. This assignment applies to the entity's publications only, and the count of citations received is not limited to the geographical collaboration status of the citing publications themselves; if an internationally collaborative publication is cited by another publication with single authorship, that citation is still counted.

This metric calculates the Citations per Publication for each type of geographical collaboration.

Collaboration Impact is a:

- Citation Impact metric

- Collaboration Metric

SciVal often displays Collaboration Impact in a chart or table with years. These years are always the years in which items were published, and do not refer to the years in which citations were received.

This metric is useful to:

- Benchmark the average scholarly impact of an entity's publications with particular types of geographical collaboration, such as:

  – The citation impact of an entity's internationally collaborative publications with that of its institutionally collaborative output.

  – The citation impact of internationally collaborative publications compared to that of non-collaborative publications with a single author.

- Compare the collaborative citation impact of entities of different sizes, but in related disciplines

- Demonstrate any benefit from establishing and maintaining collaborations to the citation impact of an entity's Scholarly Output

This metric should be used with care when:

- Comparing entities in different disciplines where the citing behavior of academics may differ:

  – Citation counts tend to be higher in disciplines such as virology, whose academics tend to publish frequently and include long reference lists, than in law, for example; these differences reflect the differences in the behavior of researchers in distinct subject fields, and not differences in performance.

  – It is not advisable to use this metric to compare entities in distinct disciplines without taking these differences into account.

  – When comparing entities made up of a mixture of disciplines, such as an Institution or Country, it is advised to apply the Subject Area filter to focus on one field that is common between all the entities.

- Understanding the citation impact of collaborative papers of small entities, when there may be gaps in their output within the Scopus coverage:

  – A single missing publication from a small data set may

significantly distort the apparent performance, whereas the buffering effect of a larger data set may compensate for 1 or 2 missing publications.

— The only way to account for this is to be vigilant, particularly when looking at small data sets such as the Scholarly Output of an early-career researcher. It is also advisable to limit the use of these metrics to comparing larger data sets in which potential gaps in the database coverage likely have a similar effect on all entities being viewed and will not invalidate the comparison.

• Entities are small so that the metric may fluctuate significantly and appear unstable over time, even when there is complete Scopus coverage:

— Collaboration Impact calculates an average value using the mean, and these types of calculations are strongly influenced by outlying publications in a small data set.

• Understanding the performance of publications in the very early stages of a new strategy, or of early-career researchers, where the short time that has passed since publication will reduce the reliability of citation information:

— These situations can be addressed by metrics that are complete immediately upon publication, such as Collaboration or Publications in Top Journal Percentiles.

Useful partner metrics are:

• **Citations per Publication**, which calculates the average citation impact across all publications within an entity regardless of their geographical collaboration status. This metric could be a useful benchmark against which to judge the extent of benefit of different types of geographical collaboration.

• **Collaboration**, which indicates the extent to which an entity's publications have international, national, or institutional co-authorship, and single authorship. This is a logical partner metric to Collaboration.

• **The set of "time-independent metrics"** which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Subject Area Count, Scopus Source Title Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

Collaboration, Collaboration Impact, Academic-Corporate Collaboration, and Academic-Corporation Collaboration Impact

5.2.3 Metric: Academic-Corporate Collaboration ❄

Academic-Corporate Collaboration in SciVal indicates the degree of collaboration between academic and corporate affiliations: to what extent are this entity's publications co-authored across the academic and corporate, or industrial, sectors?

A publication either exhibits academic-corporate collaboration, or it does not. This assignment is made based on the organization-type with which Scopus tags each affiliation.

This metric calculates the Citations per Publication for collaborative and non-collaborative papers.

Academic-Corporate Collaboration is a:

• Collaboration metric

• "Power Metric" when the "Total value" option is selected, but not when the "Percentage" option is selected

This metric is useful to:

• Investigate the degree of collaboration between the academic and corporate sectors within a data set

• Benchmark the cross-sector collaboration of entities of different sizes, but in related disciplines, such as large and small research teams, or large and small Centers of Excellence:

— It is advised to select the "Percentage" option when comparing entities of different sizes, to normalize for this variable.

• Showcase extensive collaboration between academia and industry that may underpin a set of Scholarly Output

• Investigate collaborative activity very early in a new strategy, or for an early-career researcher, since the affiliation data underlying this metric do not require time to accumulate to reliable levels in the same way as citation data do

• Look at publishing activity in a way that is difficult to manipulate

This metric should be used with care when:

• Benchmarking the extent of Academic-Corporate Collaboration of entities in different disciplines:

— The opportunity or desire to collaborate outside the sector may differ, such as between econometrics and drug discovery, or the philosophy of science and toxicology.

— It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences.

— When comparing entities made up of a mixture of disciplines, such as an Institution or Country, it is advised to apply the Subject Area filter to focus on one field that is common between all the entities.

• Entities are small and there may be gaps in their output within the Scopus coverage:

— A single missing publication from a small data set may have a significant negative impact on apparent cross-sector partnerships, whereas the effect of 1 or a few missing publication(s) from a large data set may be acceptable.

– The only way to account for this is to be vigilant, particularly when looking at small data sets such as an early-career researcher, or to limit the use of Academic-Corporate Collaboration to comparing larger data sets in which potential gaps in the database coverage likely have a similar effect on all entities being viewed and do not invalidate the comparison.

• Investigating activity in a discipline with focus outside the interest of industry, such as history:

– It is not advised to use this metric in such a situation.

Useful partner metrics are:

• Academic-Corporate Collaboration Impact, which calculates the Citations per Publication for publications with and without academic-corporate collaboration, and indicates how beneficial this cross-sector collaboration is with respect to citation impact

• The set of all other "Power Metrics" whose value tend to increase as the entity becomes bigger: Scholarly Output, Subject Area Count, Scopus Source Title Count, Citation Count, Cited Publications ("Total value"), Number of Citing Countries, Collaboration ("Total value"), Outputs in Top Citation Percentiles ("Total value"), Publications in Top Journal Percentiles ("Total value"), and h-indices

• The set of all other "time-independent metrics" which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Subject Area Count, Scopus Source Title Count, Collaboration, and Publications in Top Journal Percentiles

See Example 6, Page 62: Collaboration, Collaboration Impact, Academic-Corporate Collaboration, and Academic-Corporation Collaboration Impact

5.2.4 Metric: Academic-Corporate Collaboration Impact ❄

Academic-Corporate Collaboration Impact in SciVal indicates the citation impact of an entity's publications with or without both academic and corporate affiliations: how many citations do this entity's publications receive when they list both academic and corporate affiliations, versus when they do not?

A publication either exhibits academic-corporate collaboration, or it does not. This assignment applies to the entity's publications only, and the count of citations received is not limited to the collaboration status of the citing publications themselves; if a publication that resulted from academic-corporate collaboration is cited by another publication with only academic affiliations, that citation is still counted.

Academic-Corporate Collaboration Impact is a:

• Citation Impact metric

• Collaboration metric

SciVal often displays Academic-Corporate Collaboration Impact in a chart or table with years. These years are always the years in which

items were published, and do not refer to the years in which citations were received.

This metric is useful to:

• Benchmark the average influence of an entity's publications with and without academic-corporate collaboration, such as:

– The citation impact of publications of academic research institutes that were published in industrial collaboration, compared to those that were not

– The citation impact of publications of Researchers located in a corporate affiliation that were published in academic collaboration, compared to those that were industry-only

• Compare the citation impact of these cross-sector publications between entities of different sizes, but from related disciplines, such as large and small international networks of Researchers

• Demonstrate any benefit from establishing and maintaining academic-corporate collaborations to the citation impact of an entity's scholarly output

This metric should be used with care when:

• Comparing entities in different disciplines where the citing behavior of academics may differ:

– Citation Counts tend to be higher in disciplines such as cardiology, whose academics tend to publish frequently and include long reference lists, than in anthropology, for example; this reflects differences in the behavior of researchers in distinct subject fields, and not differences in performance.

– It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences.

– When comparing entities made up of a mixture of disciplines, such as an Institution, a Country or a Group of Countries, it is advised to apply the Research Area filter to focus on one field that is common between all the entities.

• Understanding the citation impact of collaborative papers of small entities, when there may be gaps in their output within the Scopus coverage:

– A single missing publication from a small data set may significantly distort the apparent performance, whereas the buffering effect of a larger data set may compensate for 1 or 2 missing publications.

– The only way to account for this is to be vigilant, particularly when looking at small data sets such as an early-career researcher. It is also advisable to limit the use of these metrics to comparing larger data sets in which potential gaps in the database coverage likely have a similar effect on all entities being viewed and will not invalidate the comparison.

• Entities are small so that the metric may fluctuate significantly and appear unstable over time, even when there is complete Scopus coverage:

- — Academic-Corporate Collaboration Impact calculates an average value using the mean, and these types of calculations are strongly influenced by outlying publications in a small data set.

- Understanding the performance of publications in the very early stages of a new strategy, or of early-career researchers, where the short time that has passed since publication will reduce the reliability of citation information as an input into decisions:

  - — These situations can be addressed by metrics that are useful immediately upon publication, such as Academic-Corporate Collaboration or Publications in Top Journal Percentiles.

Useful partner metrics are:

- **Citations per Publication**, that calculates the average citation impact across all publications within an entity regardless of their academic-corporate collaboration status. This metric could be a useful benchmark against which to judge the extent of benefit of different types of collaboration.

- **Academic-Corporate Collaboration**, which indicates the degree of collaboration between academic and corporate affiliations. This is a logical partner metric to Academic-Corporate Collaboration Impact.

- **The set of "time-independent metrics"** that provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Subject Area Count, Scopus Source Title Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal

Collaboration, Collaboration Impact, Academic-Corporate Collaboration, and Academic-Corporation Collaboration Impact

5.3 Published

5.3.1 Metric: Scholarly Output ❄

Scholarly Output in SciVal indicates the prolificacy of an entity: how many publications does this entity have indexed in Scopus?

Scholarly Output is a:

- Productivity metric

- Snowball Metric

- "Power Metric"

This metric is useful to:

- Benchmark the scholarly output of entities that are similar in size and that fall within similar fields of research, such as Institutions with a similar number of research staff, or Researchers in the same field of research and with similar career lengths

- Provide impressive figures to showcase the performance of entities that are large in comparison to a group of peers, when this metric is likely to give high numbers

- Look at publishing activity in a way that is difficult to manipulate

- Investigate activity during the early stages of a new strategy when publications are just starting to appear, or the scholarly output of early-career researchers

This metric should be used with care when:

- Benchmarking the productivity of entities of obviously different sizes, such as Countries and Institutions, large collaboration networks and individual Researchers, or stable, established Research Areas and small, emerging Research Areas:

  - — Differences in the value of this "Power Metric" for comparing such entities will probably reflect distinct entity sizes rather than differences in prolificacy.

  - — Users are advised to limit the use of this metric to similar entities, or to use size-normalized metrics such as Citations per Publication.

- Benchmarking the collaboration of entities with distinct disciplinary profiles:

  - — Academics in distinct fields tend to publish with very different frequencies. For example, Institutions with a high proportion of academics in life sciences, who tend to publish with high frequency, are likely to have a higher Scholarly Output than similarly sized Institutions with a high proportion of academics in humanities, who tend to publish with low frequency.

  - — Differences in Scholarly Output, in these cases, are most likely to reflect distinct disciplinary characteristics rather than to provide reliable information about differences in productivity.

  - — It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences.

  - — When comparing entities made up of a mixture of disciplines, such as an Institution or a Country, users are advised to apply the Research Area filter to focus on one field that is common between all the entities, or to use field-normalized metrics such as Publications in Top Journal Percentiles which will take this into account.

- Understanding the productivity of small entities for which there may be gaps within the Scopus coverage:

  - — A single missing publication from a small data set may have a significant negative impact on apparent productivity.

  - — The only way to account for this is to be vigilant. Consider also limiting the use of Scholarly Output to comparing larger data sets in which the potential gaps in the database coverage likely have a similar effect on all entities and do not invalidate the comparison.

Useful partner metrics are:

- **Citation Count**, which sums, over all publications, the citations received by an entity, and is a complement to Scholarly Output that counts the publications of an entity

- *h*-indices, whose values depend on a combination of Scholarly Output together with Citation Count, and are logical partner metrics

- **The set of all other "Power Metrics"** whose value tends to increase as the entity becomes bigger: Subject Area Count, Scopus Source Title Count, Citation Count, Cited Publications ("Total value"), Number of Citing Countries, Collaboration ("Total value"), Academic-Corporate Collaboration ("Total value"), Outputs in Top Citation Percentiles ("Total value"), Publications in Top Journal Percentiles ("Total value"), and *h*-indices

- **The set of all other "time-independent metrics"** which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Subject Area Count, Scopus Source Title Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

Scholarly Output, Subject Area Count and Scopus Source Title Count

5.3.2 Metric: Subject Area Count

Subject Area Count can be generated using either the Main Categories or Sub-categories. The maximum value of this metric is 27 when using Scopus Main Categories, and 334 when using Scopus Sub-categories. Publications can be assigned to a classification system in 2 ways:

"Journal-driven" assignment assumes that every publication within a journal fits within the same discipline(s) as the journal's scope. Each publication automatically adopts the subject classifications that are assigned to the journal. This method of assignment is suitable for journals that are focused in a core field, and do not tend to include publications that are also a relevant to other fields

"Publication-driven" assignment assumes that publications within a journal may have additional or different relevance to fields outside the core focus of the journal's scope. Publication-driven assignment offers the benefit of being able to assign individual publications from a journal separately to their relevant classifications. This is important for publications in multi-disciplinary journals

Subject Area Count uses "publication-driven" assignment, and a publication can be allocated to more than 1 category.

Subject Area Count is a:

- Disciplinarily metric

- "Power Metric"

This metric is useful to:

- Compare smaller entities, such as Groups of Researchers and Groups of Publication Sets, where differences in the disciplinary portfolio are most likely to be evident

- Provide impressive figures to showcase the performance of a relatively large entity, when this metric is likely to give high numbers

- Benchmark the diversity of the disciplinary portfolios of related entities, such as collaboration networks funded by the same type of grant from a given funder

- Provide evidence of cross-disciplinarily by indicating the appeal of an entity's output to readers in diverse disciplines

- Investigate differences in activity at the early stages of a new strategy when publications are just starting to appear, or between early-career researchers

- Look at publishing activity in a way that is difficult to manipulate

This metric should be used with care when:

- Comparing large entities, such as Institutions and Groups of Institutions, which will likely publish in such a large and stable journal portfolio that it will approach the maximum extent of the Scopus database so that differences may not be visible. Users are advised to narrow their focus to a slice of a large entity when using this metric, such as by applying the Research Area filter.

- Benchmarking the productivity of entities of obviously different sizes, such as departments of 150 academics with departments of 30 academics:

  – Differences in the value of this "Power Metric" when comparing such entities will probably reflect distinct entity sizes rather than differences in the diversity of their publication portfolios.

  – Users are advised to limit the use of this metric to similar entities, or to use size-normalized metrics such as Collaboration.

- Benchmarking the productivity of entities with distinct disciplinary profiles, when academics in distinct fields may have a different sized range of journals available for submission:

  – Differences in Subject Area Count, in these cases, may well reflect distinct disciplinary characteristics rather than provide reliable information about differences in publication portfolios.

  – It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences.

  – When comparing entities made up of a mixture of disciplines, such as an Institution or a Country, users are advised to apply the Research Area filter to focus on one field that is common between all the entities, or to use field-normalized metrics such as Field-Weighted Citation Impact which will take this into account.

- Understanding the productivity of small entities for which there may be gaps within the Scopus coverage

  – A single missing publication from a small data set may have a significant negative impact on the apparent breadth of a publication portfolio.

  – The only way to account for this is to be vigilant. Consider also

limiting the use of Subject Area Count to comparing slightly larger data sets in which potential gaps in the database coverage likely have a similar effect on all entities and do not invalidate the comparison.

Useful partner metrics are:

- Scopus Source Title Count, which highlights the disciplinary portfolio of an entity, and is a complement to Subject Area Count

- The set of all other "Power Metrics" whose value tends to increase as the entity becomes bigger: Scholarly Output, Scopus Source Title Count, Citation Count, Cited Publications ("Total value"), Number of Citing Countries, Collaboration ("Total value"), Academic-Corporate Collaboration ("Total value"), Outputs in Top Citation Percentiles ("Total value"), Publications in Top Journal Percentiles ("Total value"), and *h*-indices

- The set of all other "time-independent metrics" which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Scopus Source Title Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

See Example 2, Page 58: Scholarly Output, Subject Area Count and Scopus Source Title Count

5.3.3 Metric: Scopus Source Title Count

Scopus Source Title Count in SciVal indicates the diversity of an entity's disciplinary portfolio: in how many distinct titles have this entity's publications appeared?

Scopus Source Title Count can be generated using either the Main Categories or Sub-categories. The maximum value of this metric is 27 when using Scopus Main Categories, and 334 when using Scopus Sub-categories. Publications can be assigned to a classification system in 2 ways:

- **"Journal-driven" assignment** assumes that every publication within a journal fits within the same discipline(s) as the journal's scope. Each publication automatically adopts the subject classifications that are assigned to the journal. This method of assignment is suitable for journals that are focused in a core field, and do not tend to include publications that are also a relevant to other fields.

- **"Publication-driven"** assignment assumes that publications within a journal may have additional or different relevance to fields outside the core focus of the journal's scope. Publication-driven assignment offers the benefit of being able to assign individual publications from a journal separately to their relevant classifications. This is important for publications in multi-disciplinary journals.

Scopus Source Title Count uses "journal-driven" assignment, and a journal can be allocated to more than 1 category.

Scopus Source Title Count is a:

- Disciplinarity metric

- "Power Metric"

This metric is useful to:

- Compare smaller entities, such as Groups of Researchers and Groups of Publication Sets, where differences in the disciplinary portfolio are most likely to be evident

- Provide impressive figures to showcase the performance of a relatively large entity, when this metric is likely to give high numbers

- Benchmark the diversity of the disciplinary portfolios of related entities, such as collaboration networks funded by the same type of grant from a given funder

- Provide evidence of cross-disciplinarity by indicating the appeal of an entity's output to readers in diverse disciplines

- Look at publishing activity in a way that is difficult to manipulate

- Investigate differences in activity at the early stages of a new strategy when publications are just starting to appear, or between early-career researchers

This metric should be used with care when:

- Comparing large entities, such as Institutions and Groups of Countries, which will likely publish in such a large and stable disciplinary portfolio that it will approach the maximum extent of the Scopus database and differences may not be visible. Users are advised to narrow their focus to a slice of a large entity when using this metric, such as by applying the Research Area filter.

- Benchmarking the productivity of entities of obviously different sizes, such as large collaboration networks with single Publication Sets:

  - Differences in the value of this "Power Metric" when comparing such entities will probably reflect distinct entity sizes rather than differences in the diversity of their disciplinary portfolios.

  - Users are advised to limit the use of this metric to similar entities, or to use size-normalized metrics such as Academic-Corporate Collaboration.

- Benchmarking the productivity of entities with distinct disciplinary profiles, when academics in distinct fields may have different opportunities to work in a cross-disciplinary manner:

  - Differences in Scopus Source Title Count, in these cases, may well reflect distinct disciplinary characteristics rather than provide reliable information about differences in disciplinary portfolios.

  - It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences.

  - When comparing entities made up of a mixture of disciplines, such as an Institution or a Country, users are advised to apply the Research Area filter to focus on one field that is common between all the entities, or to use field-normalized metrics such as Publications in Top Journal Percentiles which will take this into account

- Understanding the productivity of small entities for which there may be gaps within the Scopus coverage:

  – A single missing publication from a small data set may have a significant negative impact on the apparent breadth of a disciplinary portfolio.

  – The only way to account for this is to be vigilant. Consider also limiting the use of Scopus Source Title Count to comparing slightly larger data sets in which potential gaps in the database coverage likely have a similar effect on all entities and do not invalidate the comparison.

Useful partner metrics are:

- Subject Area Count, which highlights the publication portfolio of an entity, and is a logical complement to Scopus Source Title Count

- The set of all other "Power Metrics" whose value tends to increase as the entity becomes bigger: Scholarly Output, Subject Area Count, Citation Count, Cited Publications ("Total value"), Number of Citing Countries, Collaboration ("Total value"), Academic-Corporate Collaboration ("Total value"), Outputs in Top Citation Percentiles ("Total value"), Publications in Top Journal Percentiles ("Total value"), and *h*-indices

- The set of all other "time-independent metrics" which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Subject Area Count, Collaboration, Academic-Corporate Collaboration, and Publication in Top Journal Percentiles

See Example 2, Page 58: Scholarly Output, Subject Area Count and Scopus Source Title Count

5.3.4 Metric: *h*-indices ❄

h-indices in SciVal indicate a balance between the productivity (Scholarly Output) and citation impact (Citation Count) of an entity's publications.

h-indices in SciVal offer 4 variants: the *h*-index, the *g*-index, the *m*-index and the *h*5-index. The *g*- and *m*-indices inherit the positive qualities of the *h*-index, but address aspects that are sometimes considered shortcomings; therefore these metrics are grouped together into a set collectively called *h*-indices:

- *h*-index is now recognized as an industry standard that gives information about the performance of Researchers and Research Areas that is very useful in some situations. *h*-index of an entity is 9 if the top 9 most-cited publications have each received at least 9 citations; it is 13 if an entity's top 13 most-cited publications have each received at least 13 citations; and so on.

- *g*-index is a variant of the *h*-index that emphasizes the most highly cited papers in a data set. The *h*-index does not give extra weighting to the most-cited publications of a data set that are likely the ones that are responsible for an entity's prestige; *g*-index can be used if this feature of the *h*-index is seen as a weakness. The *g*-index is always the same as or higher than the *h*-index.

- *m*-index is another variant of the *h*-index that displays *h*-index per year since first publication. The *h*-index tends to increase with career length, and *m*-index can be used in situations where this is a shortcoming, such as comparing researchers within a field but with very different career lengths. The *m*-index inherently assumes unbroken research activity since the first publication.

- *h*5-index uses a 5 year publication and citation window on the standard *h*-index calculation and as such can be used to fairly track the metric over time in the Benchmarking module. The *h*5-index for an entity in 2016 takes the publications published by that entity from 2012–2016 and the citations received by those publications in the same time window to form a data set. It then uses the *h*-index calculation on the data set to compute the *h*5-index.

h-indices are available in SciVal for all Researcher-based entities, and for Research Areas. The *h*5-index is also available for institution-based entities.

h-indices are:

- Productivity metrics

- Citation Impact metrics

- Snowball Metrics (*h*-index only)

- "Power Metrics"

A Researcher-based entity, or a Research Area, has a single value for the *h*-, *g*- or *m*-index that is based on all publications in the data set; these indices are not calculated per year. This is represented in the Benchmarking module as the same value for all years so that this metric can be made available in this module; in the Chart view, for instance, a horizontal line will be displayed and the years do not have a meaning.

For the *h*5-index, a single value is available for researcher-based entities and institutions. This is represented in the Benchmarking module as a trend line from 2000 until the last full year in SciVal. The publication year represents the final year of the 5 year *h*5-index, i.e. 2016 represents the year range 2012–2016 and 2015 represents the year range 2011–2015.

See Example 9, Page 67: *h*-indices

These metrics are useful to:

- Benchmark activity in a way that relies on the balance between 2 fundamental aspects of performance, namely productivity and citation impact:

  – The total number of publications, or Scholarly Output, of an entity sets a limit for the value of the h-index. If a researcher has 1 publication that has been cited 100 times, their *h*-index cannot exceed 1.

  – The total number of citations received, or Citation Count, sets the other limit for the value of the h-index. If a researcher has 100 publications which have each received 0 or 1 citations, their *h*-index also cannot exceed 1.

- Be used for a related group of metrics, each with their own strengths

- Benchmark performance over time with the h5-index as the publications and citations windows are the same length

This metric should be used with care when:

- Comparing entities of significantly different sizes:

  – The values of these metrics are limited by the Scholarly Output of an entity, and tend to increase with the size of the data set.

  – This can be accounted for within a discipline, when the difference in size is due to different career lengths, by using the m-index; in this situation, variations revealed by the m-index are due to differences in annual productivity and citations received, which are likely the performance aspects of interest.

- Benchmarking entities within different disciplines, even if these entities have similar sizes:

  – The values of *h*-indices are limited by the Citation Count of an entity, and tend to be highest in subject fields such as biochemistry, genetics and molecular biology; this reflects distinct publication and citation behavior between subject fields and does not necessarily indicate a difference in performance.

  – It is not advisable to compare the *h*-indices of entities that fall entirely into distinct disciplines, such as a Researcher in genetics with a Researcher in human-computer interaction.

  – When comparing entities made up of a mixture of disciplines, such as cross-disciplinary research teams, it is advised to apply the Research Area filter to focus on one field that is common between all the entities.

- An indication of the magnitude of the productivity and citation impact of an entity is important.

  – It is advised to use Scholarly Output and Citation Count when it is important to communicate scale.

- Entities are small and there may be gaps in their output within the Scopus coverage:

  – A single missing publication from a total or 3 or 4 will have a significant negative impact on apparent performance, whereas the effect of 1 missing publication from a total of 100 may be acceptable.

  – The only way to account for this is to be vigilant, particularly when looking at small data sets such as an early-career researcher, or to limit the use of these metrics to comparing larger data sets in which potential gaps in the database coverage likely have a similar effect on all entities being viewed and do not invalidate the comparison.

Useful partner metrics are:

- **Scholarly Output** and **Citation Count** which provide information about the magnitude of productivity and citation impact

- **The set of all other "Power Metrics"** whose value tends to increase as the entity becomes bigger: Scholarly Output, Subject Area Count, Scopus Source Title Count, Citation Count, Cited Publications ("Total value"), Number of Citing Countries, Collaboration ("Total value"), Academic-Corporate Collaboration ("Total value"), Outputs in Top Citation Percentiles ("Total value"), and Publications in Top Journal Percentiles ("Total value")

- **The set of "time-independent metrics"** that provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Subject Area Count, Scopus Source Title Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

5.4 Viewed

5.4.1 Metric: Views Count

Views Count indicates the total usage impact of an entity: how many views have this entity's publications received? Views Counts in SciVal are generated from usage data in Scopus. The metric is the sum of abstract views and clicks on the link to view the full-text at the publisher's website. These events cover all views from both subscribed and trial customers. These data hold intelligence about the interest in research outputs, and are an important piece of the jigsaw puzzle that builds the complete picture of the impact of research on academia and society.

Usage data are especially exciting for other reasons as well:

- They begin to accumulate as soon as an output is available online, and are more immediate than citation activity, so that an emerging trend or research talent may be more quickly spotted than via citation activity.

- They reflect the interest of the whole research community, including undergraduate and graduate students, and researchers operating in the corporate sector, who tend not to publish and cite and who are "hidden" from citation-based metrics.

- They can help to demonstrate the impact of research that is published with the expectation of being read rather than extensively cited, such as clinical and arts and humanities research.

- Scopus usage data are COUNTER-compliant, and are audited every year. COUNTER (Counting Online Usage of Networked Electronic Resources) is an international initiative serving librarians, publishers and intermediaries by setting standards that facilitate the recording and reporting of online usage statistics in a consistent, credible and compatible way.

- The Views Count metrics are anonymized. They do not provide any information about what users of a particular institution are viewing, and it is not possible to see the usage of a particular customer.

Views Count is a:

- Usage Impact metric

- "Power Metric": its value tends to increase as the size of the entity increases

Views Count may be displayed in a chart or table with months and/or years:

- In SciVal, the year shows the date on which items became available online; this may be different to the official publication date. They do not refer to the years in which publications were viewed

This metric is useful to:

- Benchmark the views received by entities of similar size, and that fall into similar disciplines, such as multidisciplinary institutions with a similar number of research staff, or international collaboration networks in similar disciplines

- Showcase the performance of entities that are large in comparison to a group of peers, when this metric is likely to give high numbers

- Showcase the performance of entities that have published a few noticeably highly viewed publications that will have a positive effect on the total for the entire data set

- Give an early indication of interest in output that has recently become available, for example in the very early stages of a new strategy, or of early-career researchers

- Showcase the interest of the whole research community, and not only the two-thirds who publish and therefore cite. The one-third which does not tend to publish includes large numbers of under-graduate and graduate students, as well as researchers operating in the corporate sector

- Demonstrate interest in outputs produced in disciplines with low citation potential, such as clinical research and the arts and humanities that are generally well read but poorly cited

This metric should be used with care when:

- Benchmarking the visibility of entities of obviously different sizes, when this "Power Metric" may most closely reflect entity size rather than differences in views received. Users are advised to use the size-normalized metrics Views per Publication or Field-Weighted Views Impact to compare the visibility of entities of different sizes.

- Benchmarking the usage of entities with distinct disciplinary profiles. The average usage between disciplines is variable and it is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences. When comparing entities made up of a mixture of disciplines, such as an Institution or an interdisciplinary research group, it is advised to apply a Research Area filter to focus on one field that is common between all the entities, or to select Field-Weighted Views Impact which will take this into account.

- Revealing the extent to which each of an entity's outputs are viewed, since one or a few publications with a very high number of views can

conceal a sizeable body of unviewed or poorly viewed material

- There may be gaps in output in the database coverage:

  - For Scopus usage, this will mainly apply when entities are small, and a single missing output may have a significant negative impact on apparent usage.

  - The only way to account for this is to be vigilant; consider also limiting the use of Views Count to comparing larger data sets in the same discipline in which potential gaps in the database coverage likely have a similar effect on all entities being viewed and do not invalidate the comparison.

- The people who will use the metrics do not like to see a trend that "dips" in recent years:

  - This typically happens with Views Count because the most recent outputs have had less time to receive views than older ones. Users are advised to use Field-Weighted Views Impact to avoid this drop, if it is of concern.

Useful partner metrics are:

- **Views per Publication** and **Field-Weighted Views Impact**, which bring complementary perspectives on total views received. Both account for differences in the size of entities being compared, and Field-Weighted Views Impact also accounts for differences in viewing behavior between disciplines.

- **Field-Weighted Views Impact** avoids the "dip" in recent years due to the most recent outputs having had less time to receive views than older ones

5.4.2 Metric: Outputs in Top Views Percentiles

Outputs in Top Views Percentiles in SciVal indicates the extent to which an entity's publications are present in the most-viewed percentiles of a data universe: how many publications are in the top 1%, 5%, 10% or 25% of the most-viewed publications?

The entire Scopus database, or "World", is the default data universe used to generate this metric:

- The view counts that represent the thresholds of the 1%, 5%, 10% and 25% most-viewed papers in Scopus per Publication Year are calculated. Sometimes the same number of views received by the publication at, say, the 10% boundary has been received by more than 10% of publications; in this case, all of the publications that have received this number of views are counted within the top 10% of the Scopus data universe, even though that represents more than 10% by volume.

- SciVal uses these view thresholds to calculate the number of an entity's publications that fall within each percentile range.

Use of the Publication Type filter affects the publications in the data universe that are used to generate the view thresholds, as well as the publications of the entity upon which the calculation is performed.

Outputs in Top Views Percentiles can only be calculated for the

current year from the first data snapshot on or after mid-July. It will be displayed as a null value until this date is reached. This metric depends on being able to divide the publications into 100 percentiles, and this level of division is not possible earlier in the publication year when items just published have received very few citations.

When field-weighting Outputs in Top Views Percentiles, the document views ratio is used instead of views to compute values for each percentile. The rest of the calculation remains the same as the calculation for the Field-Weighted Citation Impact. The value displayed in the chart or table view is the number of outputs that meet the benchmark, so contrary to the Field-Weighted Citation Impact, the field-weighted value for Outputs in Top Views Percentiles will not be near 1.00.

This metric is useful to:

- Benchmark the contributions towards the most influential, highly viewed publications in the world of entities of different sizes, but in similar disciplines

- It is advised to select the "Percentage" option when comparing entities of different sizes, to normalize for this variable

- Distinguish between entities whose performance seems similar when viewed by other metrics, such as Scholarly Output, Views per Publication, or Collaboration

- Showcase the performance of a prestigious entity whose publications are amongst the most viewed and highly visible publications of the scholarly world

This metric should be used with care when:

- Comparing entities in different disciplines:

  – It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences.

  – When comparing entities made up of a mixture of disciplines, such as an Institution or Country, it is advised to apply the Research Area filter to focus on one field that is common between all the entities.

- Trust needs to be built in the metrics in SciVal. The views thresholds may depend on the entire Scopus database, and it will be difficult for a user to validate these boundaries. Users are advised to select simpler metrics, such as Views per Publication, if trust in the accuracy of the SciVal calculations needs to be built.

Useful partner metrics are:

- **Views per Publication,** which indicates the average usage impact of an entity's publications

- **The set of all other "Power Metrics"** whose value tends to increase as the entity becomes bigger: Scholarly Output, Subject Area Count, Scopus Source Title Count, Citation Count, Cited Publications ("Total value"), Number of Citing Countries, Collaboration ("Total value"), Academic-Corporate Collaboration ("Total value"), Publications in Top Journal Percentiles ("Total value"), and h-indices

- **The set of "time-independent metrics"** which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Subject Area Count, Scopus Source Title Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

5.4.3 Metric: Views per Publication

Views per Publication indicates the average usage impact of an entity's publications: how many views have this entity's publications received on average? Views per Publication is calculated analogously to Citations per Publication.

Views per Publication is a:

- Usage Impact metric

Views per Publication may be displayed in a chart or table with months and/or years:

- In SciVal, the year used is based on the date on which items became available online; this may be different to the official publication date. They do not refer to the years in which publications were viewed

This metric is useful to:

- Benchmark the average usage impact of publications within a body of work or entity

- Compare the average visibility of publications of entities of different sizes, but in related disciplines, such as Researchers working in a similar Research Area

- Showcase the performance of entities that have published a few highly viewed papers that will have a positive effect on the average of the entire data set

- Give an early indication of interest in output that has recently become available, for example in the very early stages of a new strategy, or of early-career researchers

- Showcase the engagement of the whole research community, and not only the two-thirds who publish and therefore cite. The one-third which does not tend to publish includes large numbers of undergraduate and graduate students, as well as researchers operating in the corporate sector.

- Demonstrate interest in output produced in disciplines with low citation potential, such as clinical research and the arts and humanities that are generally well read but poorly cited

This metric should be used with care when:

- Benchmarking the usage of entities with distinct disciplinary profiles:

  – The average usage between disciplines is variable and it is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences.

– When comparing entities made up of a mixture of disciplines, such as an interdisciplinary collaboration network, it is advised to apply a Research Area filter to focus on one field that is common between all the entities, or to select Field-Weighted Views Impact which will take this into account.

• Revealing the extent to which each of an entity's outputs are viewed, since one or a few publications with a very high number of views can conceal a sizeable body of unviewed or poorly viewed material

• There may be gaps in output in the database coverage:

– For Scopus usage, this will mainly apply when entities are small, and a single missing publication may have a significant negative impact on apparent usage.

– The only way to account for this is to be vigilant; consider also limiting the use of Views per Publication to comparing larger data sets in the same discipline where gaps in the database coverage likely have a similar effect on all entities being viewed and do not invalidate the comparison.

• Entities are small, such that the metric may fluctuate significantly and appear unstable over time, even when there is complete database coverage. Views per Publication calculates an average value, and is strongly influenced by outlying publications in a small data set.

• The people who will use the metrics do not like to see a trend that "dips" in recent years:

– This typically happens with Views per Publication because the most recent publications have had less time to receive views than older ones. Users are advised to use Field-Weighted Views Impact to avoid this drop, if it is of concern.

Useful partner metrics are:

• **Field-Weighted Views Impact**, which is a complement to Views per Publication and takes into account behavioral differences between disciplines. Field-Weighted Views Impact also avoids the "dip" in recent years due to the most recent publications having had less time to receive views than older ones.

5.4.4 Metric: Field-Weighted Views Impact

Field-Weighted Views Impact indicates how the number of views received by an entity's publications compares with the average number of views received by all other similar publications in the same data universe: how do the views received by this entity's publications compare with the world average for that database?

Similar publications are those publications in the database that have the same publication year, publication type, and discipline, as represented by the Scopus classification system.

• A Field-Weighted Views Impact of 1.00 indicates that the entity's publications have been viewed exactly as would be expected based on the global average for similar publications in the same database; the Field-Weighted Views Impact of "World", that is, the entire Scopus database, is 1.00.

• A Field-Weighted Views Impact of more than 1.00 indicates that the entity's publications have been viewed more than would be expected based on the global average for similar publications in the same database; for example, 3.87 means 287% more views than world average within the same database.

• A Field-Weighted Views Impact of less than 1.00 indicates that the entity's publications have been viewed less than would be expected based on the global average for similar publications in the same database; for example, 0.55 means 45% fewer views than world average within the same database.

• Publications can be allocated to more than one category in the Scopus classification system. When we calculate the expected views for similar publications, it is important that these multi-category publications do not exert too much weight; for example, if a publication P belongs to both parasitology and microbiology, it should not have double the influence of a publication that belongs to only one or the other of these.

This is accounted for in this metric calculation by distributing publication and views counts equally across multiple categories; publication P would be counted as 0.5 publications for each of parasitology and microbiology, and its views would also be shared equally between them.

Field-Weighted Views Impact is a:

• Usage Impact metric

Field-Weighted Views Impact may be displayed in a chart or table with months and/or years:

• In SciVal, the year shows the date on which items became available online, this may be different to the official publication date. They do not refer to the years in which publications were viewed

This metric is useful to:

• Benchmark entities regardless of differences in their size, disciplinary profile, age, and publication-type composition, such as an institution and departments within that institution

• Easily understand the prestige of an entity's usage performance by observing the extent to which its Field-Weighted Views Impact is above or below the world average of 1.00

• Present usage data in a way that inherently takes into account the lower number of views received by relatively recent publications, thus avoiding the dip in recent years seen with Views Count and Views per Publication

• Gain insight into the usage performance of an entity in a discipline with relatively poor database coverage, since gaps in the database will apply equally to the entity's publications and to the set of similar publications

• Use as a default to view usage data, since it considers multiple variables that can affect other metrics

• Give an early indication of the interest in output that has recently become available, for example in the very early stages of a new strategy, or of early-career researchers

- Showcase the engagement of the whole research community, and not only of the two-thirds who publish and therefore cite. The one-third which does not tend to publish includes large numbers of undergraduate and graduate students, as well as researchers operating in the corporate sector.

- Demonstrate interest in output produced in disciplines with low citation potential, such as clinical research and the arts and humanities that are generally well read but poorly cited

This metric should be used with care when:

- Information about the magnitude of the number of views received by an entity's publications is important. In these situations, it is advised to use Views Count or Views per Publication.

- Demonstrating excellent performance to those who prefer to see high numbers; Views Count or Views per Publication would be more suitable in these circumstances

- Entities are small, such that the metric may fluctuate significantly and appear unstable over time, even when there is complete database coverage. Field-Weighted Views Impact calculates an average value, and is strongly influenced by outlying publications in a small data set.

- Trust needs to be built in research metrics. This calculation accounts for multiple normalizations, and the generation of the average views for similar publications requires calculations on the entire database which will be difficult for a user to validate. Users are advised to select simpler metrics, such as Views Count or Views per Publication, if trust in the accuracy of the metrics calculations needs to be built.

- Completely answering every question about performance from a usage perspective. Field-Weighted Views Impact is a very useful metric and accounts for several variables, but using it to the exclusion of other metrics severely restricts the richness and reliability of information that a user can draw on.

Useful partner metrics are:

- **Views Count** and **Views per Publication**. They indicate the magnitude of the number of views received, to complement the relative view offered by Field-Weighted Views Impact. They are also simple and allow transparency on the underlying data to build trust in the accuracy of metric calculations.

5.5 Cited

5.5.1 Metric: Citation Count ❄

Citation Count in SciVal indicates the total citation impact of an entity: how many citations have this entity's publications received?

Citation Count is a:

- Citation Impact metric

- Snowball Metric

- "Power Metric"

SciVal often displays Citation Count in a chart or table with years. These years are always the years in which items were published, and do not refer to the years in which citations were received.

This metric is useful to:

- Benchmark the visibility of entities of similar size, and that fall into similar disciplines, such as multidisciplinary institutions with a similar number of research staff, or international collaboration networks in similar disciplines

- Provide impressive figures to showcase the performance of entities that are large in comparison to a group of peers, when this metric is likely to give high numbers

- Showcase the performance of entities that have published a few noticeably highly cited papers

This metric should be used with care when:

- Benchmarking the visibility of entities of obviously different sizes, when this "Power Metric" will most closely reflect entity size rather than differences in visibility; a Group of Institutions such as a US state will generally have a higher Citation Count than an individual institution, for example. Users are advised to use the size-normalized Citation Impact metrics Citations per Publication or Field-Weighted Citation Impact to compare the visibility of entities of different sizes.

- Benchmarking the collaboration of entities with distinct disciplinary profiles:

  – Academics working in medical sciences and in virology, for instance, tend to publish and cite with high frequencies, whereas those working in business schools or in linguistics tend to publish and cite with lower frequencies.

  – It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences.

  – When comparing entities made up of a mixture of disciplines, such as an Institution or a Country, it is advised to apply the Research Area filter to focus on one field that is common between all the entities, or to select Field-Weighted Citation Impact which will take this into account.

- Investigating the reliability with which an entity's publications are cited, since one or a few publications with a very high number of citations can conceal a sizeable body of uncited material. Users are advised to select Cited Publications to give an idea of reliability.

- Entities are small and there may be gaps in their output within the Scopus coverage:

  – A single missing publication from a small data set may have a significant negative impact on apparent visibility. For example, an academic's performance may suffer due to gaps in items they have published, as well as gaps in publications citing the publications that are indexed.

ELSEVIER

— The only way to account for this is to be vigilant. Consider also limiting the use of Citation Count to comparing larger data sets in the same discipline in which potential gaps in the database coverage likely have a similar effect on all entities being viewed and do not invalidate the comparison.

• There is a concern that excessive self-citations may be artificially inflating the number of citations. Users can judge whether the level of self-citations is higher than normal by deselecting the "Include self-citations" option.

• Uncovering the performance of publications in the very early stages of a new strategy, or of early- career researchers, where the short time that has passed since publication will reduce the reliability of basing decisions on citation information. Users are advised to use metrics such as Scholarly Output or Collaboration in these situations.

• The person who will use the data does not like to see a line that "dips" in recent years. This typically happens with Citation Count because recent publications have had little time to receive citations. Users are advised to use Field-Weighted Citation Impact or Publications in Top Journal Percentiles to avoid this drop, if it is of concern.

Useful partner metrics are:

• **Citations per Publication** and **Field-Weighted Citation Impact**, which bring complementary perspectives into view on total visibility, and also account for differences in the size of entities being compared. Field-Weighted Citation Impact also accounts for differences in publication and citation behavior between disciplines.

• **Field-Weighted Citation Impact**, **Outputs in Top Citation Percentiles**, and **Publications in Top Journal Percentiles** avoid the "dip" in recent years due to little time having passed to receive citations since publication

• **Cited Publications** provides a measure of the reliability that an entity's publications will subsequently be cited, and is not affected by a high number of citations received by 1 or a few publications

• *h-indices*, whose values depend on a combination of Citation Count together with Scholarly Output, and are logical partner metrics

• **The set of all other "Power Metrics"** whose value tends to increase as the entity becomes bigger: Scholarly Output, Subject Area Count, Scopus Source Title Count, Cited Publications ("Total value"), Number of Citing Countries, Collaboration ("Total value"), Academic-Corporate Collaboration ("Total value"), Outputs in Top Citation Percentiles ("Total value"), Publications in Top Journal Percentiles ("Total value"), and h-indices

• **The set of "time-independent metrics"** that provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Subject Area Count, Scopus Source Title Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

5.5.2 Metric: Field-Weighted Citation Impact ❄

Field-Weighted Citation Impact in SciVal indicates how the number of citations received by an entity's publications compares with the average number of citations received by all other similar publications in the data universe: how do the citations received by this entity's publications compare with the world average?

• A Field-Weighted Citation Impact of 1.00 indicates that the entity's publications have been cited exactly as would be expected based on the global average for similar publications; the Field-Weighted Citation Impact of "World", or the entire Scopus database, is 1.00.

• A Field-Weighted Citation Impact of more than 1.00 indicates that the entity's publications have been cited more than would be expected based on the global average for similar publications; for example, 2.11 means 111% more than the world average.

• A Field-Weighted Citation Impact of less than 1.00 indicates that the entity's publications have been cited less than would be expected based on the global average for similar publications; for example, 0.87 means 13% less than the world average.

Similar publications are those publications in the Scopus database that have the same publication year, publication type, and discipline, as represented by the Scopus journal classification system:

• Publications can be assigned to a classification system in 2 ways:

— **Journal-driven" assignment** assumes that every publication within a journal fits within the same discipline(s) as the journal's scope. Each publication automatically adopts the subject classifications that are assigned to the journal. This method of assignment is suitable for journals that are focused in a core field, and do not tend to include publications that are also relevant to other fields.

— **"Publication-driven" assignment** assumes that publications within a journal may have additional or different relevance to fields outside the core focus of the journal's scope. Publication-driven assignment offers the benefit of being able to assign individual publications from a journal separately to their relevant classifications. This is important for publications in multi-disciplinary journals.

• Field-Weighted Citation Impact uses "publication-driven" assignment

• Publications are allocated to the classification Sub-category level, and can be allocated to more than one Sub-category. When calculating the expected citations for similar publications, it is important that these multi-category publications do not exert too much weight; for example, if a publication P belongs to both in both parasitology and microbiology, it should not have double the influence of a publication that belongs to only one or the other Sub-category. This is accounted for in SciVal by distributing publication and citation counts equally

across multiple journal categories; publication P would be counted as 0.5 publications for each of parasitology and microbiology, and its citations would be shared equally between these Sub-categories.

Field-Weighted Citation Impact is a:

- Citation Impact metric
- Snowball Metric

SciVal often displays Field-Weighted Citation Impact in a chart or table with years. These years are always the years in which items were published, and do not refer to the years in which citations were received. The citations received in the year in which an item was published, and the following 3 years, are counted for this metric

This metric is useful to:

- Benchmark entities regardless of differences in their size, disciplinary profile, age, and publication-type composition, such as.

  – An institution and departments (Groups of Researchers) within that institution

  – A country and small research institutes within that country

  – A geographical region and countries within that region

- Easily understand the prestige of an entity's citation performance by observing the extent to which its Field-Weighted Citation Impact is above or below the world average of 1.00

- Present citation data in a way that inherently takes into account the lower number of citations received by relatively recent publications, thus avoiding the dip in recent years seen with Citation Count and Citations per Publication

- Gain insight into the relative citation performance of an entity in a discipline with relatively poor Scopus coverage, since gaps in the database will apply equally to the entity's publications and to the set of similar publications

- Use as a default to view citation data, since it takes multiple variables that can affect other metrics into account

- Look at publishing activity in a way that is difficult to manipulate

This metric should be used with care when:

- Information about the magnitude of the number of citations received by an entity's publications is important. In these situations, it is advised to use Citation Count or Citations per Publication.

- Demonstrating excellent performance to those who prefer to see high numbers; Citation Count or Citations per Publication would be more suitable in these circumstances

- Entities are small so that the metric may fluctuate significantly and appear unstable over time, even when there is complete Scopus coverage. Field-Weighted Citation Impact calculates an average value, and these types of calculations are strongly influenced by outlying publications in a small data set. For example one or two highly cited articles will have a much larger effect on an entity made

up of 10 articles, than an entity consisting of 1,000 articles, which can lead to an inflated value.

- Entities contain a large proportion of recently published outputs. Both the output universe (publications in the same year, subject area and document type) and the citation universe (publications that are citing the publication universe), can affect FWCI fluctuations which may be seen in the metric values in the period immediately following date of publication. For an output published in the current year, new outputs will be added to the Scopus database during the remainder of the year and so the publication universe for the metric is still being defined. With the citation universe, as citations take time to accrue, this can also lead to greater fluctuations in the initial period after publication. Over time both of these effects will reduce, so smaller fluctuations in the metric should be seen.

- Trust needs to be built in the metrics in SciVal. This calculation accounts for multiple normalizations, and the generation of the average citations for similar publications requires a view on the entire Scopus database which will be difficult for a user to validate. Users are advised to select simpler metrics, such as Citation Count or Citations per Publication, if trust in the accuracy of the SciVal calculations needs to be built.

- Uncovering the performance of publications in the very early stages of a new strategy, or of early-career researchers, where the short time that has passed since publication will reduce the reliability of basing decisions on citation information. It is advised to use on of the time-independent metrics, such as Academic-Corporate Collaboration, in these cases.

- Completely answering every question about performance. Field-Weighted Citation Impact is a very useful metric, but using it to the exclusion of other metrics severely restricts the richness and reliability of information that a user can draw from SciVal

Useful partner metrics are:

- **Citation Count** and **Citations per Publication**. They indicate the magnitude of the number of citations received, to complement this relative view offered by Field-Weighted Citation Impact. They are also simple and offer transparency on the underlying data to build trust in SciVal's metrics calculations.

- **The set of "time-independent metrics"** that provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Subject Area Count, Scopus Source Title Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

Mathematical notation

The Field-Weighted Citation Impact (FWCI) for a set of N publications is defined as:

$$FWCI \equiv \frac{1}{N} \sum_{i=1}^{N} \frac{c_i}{e_i}$$

c_i = citations received by publication i

e_i = expected number of citations received by all similar publications in the publication year plus following 3 years

When a similar publication is allocated to more than 1 discipline, the harmonic mean is used to calculate e_i. For a publication i that is part of 2 disciplines:

$$\frac{1}{e_i} = \frac{1}{e_i}\left(\frac{1}{e_A} + \frac{1}{e_B}\right)$$

e_A, e_B = fractional counts of publications and citations, so that publication i will be counted as 0.5 publications in each of eA and eB, and the citations it has received will also be shared between A and B

See Example 5, Page 61: Field-Weighted Citation Impact

5.5.3 Metric: Outputs in Top Citation Percentiles ❄

Outputs in Top Citation Percentiles in SciVal indicates the extent to which an entity's publications are present in the most-cited percentiles of a data universe: how many publications are in the top 1%, 5%, 10% or 25% of the most-cited publications?

The entire Scopus database, or "World", is the data universe used to generate this metric:

- The citation counts that represent the thresholds of the 1%, 5%, 10% and 25% most-cited papers in Scopus per Publication Year are calculated. Sometimes the same number of citations received by the publication at, say, the 10% boundary has been received by more than 10% of publications; in this case, all of the publications that have received this number of citations are counted within the top 10% of the Scopus data universe, even though that represents more than 10% by volume.

- SciVal uses these citation thresholds to calculate the number of an entity's publications that fall within each percentile range.

Use of the Publication Type filter affects the publications in the data universe that are used to generate the citation thresholds, as well as the publications of the entity upon which the calculation is performed. The exclusion of self-citations affects only the entity, and not the data universe.

Outputs in Top Citation Percentiles can only be calculated for the current year from the first data snapshot in early June. It will be displayed as a null value until this date is reached. This metric depends on being able to divide the publications into 100 percentiles, and this level of division is not possible earlier in the publication year when items just published have received very few citations.

When field weighting Outputs in Top Citation Percentiles, the document citation ratio is used instead of citations to compute values for each percentile. The document citation ratio is the number of citations the document has received divided by the expected number of citations. The rest of the calculation remains the same as the calculation for the Field-Weighted Citation Impact. The value displayed in the chart or table view is the number of outputs that meet the benchmark, so contrary to the Field-Weighted Citation Impact, the field-weighted value for Outputs in Top Citation Percentiles will not be near 1.00.

Outputs in Top Citation Percentiles is a:

- Citation Impact metric

- Snowball Metric

- "Power Metric" when the "Total value" option is selected, but not when the "Percentage" option is selected

SciVal often displays Outputs in Top Citation Percentiles in a chart or table with years. These years are always the years in which items were published, and do not refer to the years in which citations were received.

This metric is useful to:

- Benchmark the contributions towards the most influential, highly cited publications in the world of entities of different sizes, but in similar disciplines

- It is advised to select the "Percentage" option when comparing entities of different sizes, to normalize for this variable

- Distinguish between entities whose performance seems similar when viewed by other metrics, such as Scholarly Output, Citations per Publication, or Collaboration

- Showcase the performance of a prestigious entity whose publications are amongst the most cited and highly visible publications of the scholarly world

- Present citation data in a way that inherently considers the lower number of citations received by relatively recent publications, thus avoiding the dip in recent years seen with Citation Count and Citations per Publication

This metric should be used with care when:

- Comparing entities in different disciplines:

  - Citation counts tend to be higher in disciplines such as immunology and microbiology, whose academics tend to publish frequently and include long reference lists, than in mathematics, where publishing 1 item every 5 years that refers to 1 or 2 other publications is common; these differences reflect the distinct behavior of researchers in distinct subject fields, and not differences in performance.

ELSEVIER

— It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences.

— When comparing entities made up of a mixture of disciplines, such as an Institution or Country, it is advised to apply the Research Area filter to focus on one field that is common between all the entities.

• Entities are small and there may be gaps in their output within the Scopus coverage. A single missing highly cited publication from a small data set will have a significant negative impact on apparent performance. Although it is relatively unlikely that such prominent publications are not indexed by Scopus, we advise users to be vigilant and to bear this possible limitation in mind.

• There is a concern that excessive self-citations may be artificially inflating the number of publications that appear in the top percentiles. Users can judge whether the level of self-citations is higher than expected by deselecting the "Include self-citations" option.

• Understanding the status of publications of an early-career researcher, or those resulting from the first stages of a new strategy, where insufficient time may have passed to ensure that presence in top citation percentiles is a reliable indicator of performance. These situations can be addressed by metrics that are useful immediately upon publication, such as Publications in Top Journal Percentiles.

• Trust needs to be built in the metrics in SciVal. The citation thresholds may depend on the entire Scopus database, and it will be difficult for a user to validate these boundaries. Users are advised to select simpler metrics, such as Citations per Publication, if trust in the accuracy of the SciVal calculations needs to be built.

Useful partner metrics are:

• **Cited Publications**, which indicates the reliability with which an entity's output is built on by subsequent research by counting publications that have received at least 1 citation. It is not affected by 1 or a few very highly cited publications.

• **Publications in Top Journal Percentiles**, which indicates the extent to which an entity's publications are present in the most-cited journals in the data universe, and is independent of the citations received by the publications themselves. This is a logical partner metric.

• **The set of all other "Power Metrics"** whose value tends to increase as the entity becomes bigger: Scholarly Output, Subject Area Count, Scopus Source Title Count, Citation Count, Cited Publications ("Total value"), Number of Citing Countries, Collaboration ("Total value"), Academic-Corporate Collaboration ("Total value"), Publications in Top Journal Percentiles ("Total value"), and *h*-indices

• **The set of "time-independent metrics"** which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Subject Area Count, Scopus Source Title Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

Outputs in Top Citation Percentiles

5.5.4 Metric: Publications in Top Journal Percentiles ❄

Publications in Top Journal Percentiles in SciVal indicates the extent to which an entity's publications are present in the most-cited journals in the data universe: how many publications are in the top 1%, 5%, 10% or 25% of the most-cited journals indexed by Scopus?

The most-cited journals are defined by the journal metrics CiteScore, SNIP (Source-Normalized Impact per Paper) or SJR (SCImago Journal Rank). This means that the data universe is the set of items indexed by Scopus that have a journal metric and so can be organized into percentiles; this excludes publications in stand-alone books and trade publications, which do not have journal metrics.

All items indexed by Scopus that have a CiteScore, SNIP or SJR value form the data universe used to generate this metric:

• The CiteScore, SNIP or SJR values at the thresholds of the top 1%, 5%, 10% and 25% most-cited journals in Scopus per Publication Year are calculated. It is possible that the same journal metric value received by the indexed item at, say, the 10% boundary has been received by more than 10% of indexed items; in this case, all items with this journal metric value are counted within the top 10% of the data universe, even though that represents more than 10% items by volume. This is less likely to happen than for the Outputs in Top Citation Percentiles thresholds, because the journal metrics are generated to 2 or 3 decimal places which reduces the chance of items having the same value.

• These thresholds are calculated separately for CiteScore, SNIP and SJR, not once for a combination of both journal metrics.

• For CiteScore, the percentage thresholds are taken directly from the CiteScore Percentile values that are calculated by Scopus. A journal receives a CiteScore Percentile for each ASJC in which it's categorized. SciVal always uses the highest relevant CiteScore Percentile, which is dictated by the subject area filter.

• SciVal uses these journal metric thresholds to calculate the number of an entity's publications within indexed items that fall within each percentile range.

• Indexed items have multiple journal metric values, for distinct years. Which one is used in assigning publications to journal metrics thresholds?

— The journal metric value matching the publication year of an item of scholarly output is used as far as possible.

— The first journal metric values for SNIP and SJR are available for 1999. For scholarly output published in the range 1996-1999, the journal metric value for 1999 is used.

ELSEVIER

- The first journal metric values for CiteScore are available for 2011. For scholarly output published in the range 1996–2010, the CiteScore Percentile value for 2011 is used. If no 2011 metrics are available for the journal, then no value is displayed.

- Current year journal metric values are published during the course of the following year. For recent items whose journal metric values have not yet been published, the previous year's journal metric values are used until the current year's become available.

- A publication may be counted in the Top Journal Percentiles without itself ever having been cited. The citations received by an individual publication are irrelevant for this metric, which is based only on citations received by a journal or conference proceedings.

SNIP and SJR are both field-normalized journal metrics, meaning that this metric can be used to compare the presence of publications in journals of entities working in different disciplines. The CiteScore Percentile allows a similar comparison across disciplines.

Publications in Top Journal Percentiles is a:

- Citation Impact metric

- "Power Metric" when the "Total value" option is selected, but not when the "Percentage" option is selected

SciVal often displays Publications in Top Journal Percentiles in a chart or table with years. These years are always the years in which items were published, and do not refer to the years in which citations were received.

This metric is useful to:

- Benchmark entities even if they have different sizes and disciplinary profiles:

  - It is advised to select the "Percentage" option when comparing entities of different sizes, to normalize for this variable.

  - SNIP and SJR are field-normalized metrics. CiteScore is not a field-normalized metric, but the CiteScore Percentile (used for the calculation of the Publications in Top Journal Percentiles) is. Selecting one of these journal metrics will inherently account for differences in the behavior of academics between fields.

- Showcase the presence of publications in journals that are likely to be perceived as the most prestigious in the world

- Incorporate peer review into a metric, since it is the judgment of experts in the field that determines whether a publication is accepted by a particular journal or not

- Avoid the "dip" in recent years seen with metrics like Citation Count and Citations per Publication

- Investigate performance very early in a new strategy, or for an early-career researcher, since journal data underpin this metric, and not the citations received by individual publications themselves which would require time to accumulate

- The journal metrics for the entire database are available for download in a spreadsheet from https://www.elsevier.com/solutions/scopus/features/metrics, so that a user can determine the thresholds themselves if required.

This metric should be used with care when:

- The objective is to judge an entity's publications based on their actual performance, rather than based on the average of a journal:

  - A publication may appear in a journal with a very high CiteScore, SNIP or SJR value, and not itself receive any citations; even the most highly cited journals in the world contain publications that have never been cited.

  - A publication may be very influential and have received many citations, without being published in a journal with a high CiteScore, SNIP or SJR value; journals which do not rank very highly in the data universe may still contain very highly cited publications.

- Benchmarking the performance of a Researcher who is the editor of a journal in the top percentiles, since they can publish multiple editorials which all fall into these top percentiles. In this situation, it is advised to use the Publication Type filter to exclude editorials and ensure that the types of publications in the data sets being compared are consistent.

- Entities are small and there may be gaps in their output within the Scopus coverage. A single missing publication from a small data set will have a significant negative impact on apparent performance. Although it is relatively unlikely that publications in such prominent journals are not indexed by Scopus, we advise users to be vigilant and to bear this possible limitation in mind.

Useful partner metrics are:

- **Citation Count**, **Citations per Publication**, **Field-Weighted Citation Impact**, and **Outputs in Top Citation Percentiles** which rely on the citations received by an entity's publications themselves, and not on the average performance of the journal

- **Outputs in Top Citation Percentiles**, which indicates the extent to which an entity's publications are present in the most-cited percentiles of the data universe, but depends on the citations received by the publications themselves. This is a logical partner metric.

- **The set of all other "Power Metrics"** whose value tends to increase as the entity becomes bigger: Scholarly Output, Subject Area Count, Scopus Source Title Count, Citation Count, Cited Publications ("Total value"), Number of Citing Countries, Collaboration ("Total value"), Academic-Corporate Collaboration ("Total value"), Outputs in Top Citation Percentiles ("Total value"), and *h*-indices

- **The set of all other "time-independent metrics"** which provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Subject Area Count, Scopus Source Title Count, Collaboration, and Academic-Corporate Collaboration

5.5.5 Metric: Citations per Publication ❄

Citations per Publication in SciVal indicates the average citation impact of each of an entity's publications: how many citations have this entity's publications received on average?

Citations per Publication is a:

- Citation Impact metric

- Snowball Metric

SciVal often displays Citations per Publication in a chart or table with years. These years are always the years in which items were published, and do not refer to the years in which citations were received.

This metric is useful to:

- Benchmark the average citation impact of publications within a body of work

- Compare the average influence of publications of entities of different sizes, but in related disciplines, such as research teams working in a similar field of research, or a Researcher and Publication Sets belonging to that Researcher

- Showcase the performance of entities that have published a few notably highly cited papers that will have a positive effect on the average of the entire data set

This metric should be used with care when:

- Benchmarking the average influence of the publications of entities with distinct disciplinary profiles, such as Institutions with sizeable humanities schools with Institutions without humanities schools

    – It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences.

    – When comparing entities made up of a mixture of disciplines, such as an Institution or a Country, it is advised to apply the Research Area filter to focus on one field that is common between all the entities, or to select Field-Weighted Citation Impact which will account for disciplinary differences.

- Investigating the reliability with which an entity's publications are cited, since one or a few publications with a very high number of citations can conceal a sizeable body of uncited material. Users are advised to use Cited Publications to investigate the proportion of publications in a data set that have been cited.

- Entities are small and there may be gaps in their output within the Scopus coverage:

    – A single missing publication from a small data set may have a significant negative impact on apparent visibility. For example, an academic's performance may suffer due to gaps in items they have published, as well as gaps in items citing the publications that are indexed.

    – The only way to account for this is to be vigilant. Consider also limiting the use of Citations per Publication to comparing larger data sets in the same discipline in which potential gaps in the database coverage likely have a similar effect on all entities being viewed and do not invalidate the comparison.

- Entities are small so that the metric may fluctuate significantly and appear unstable over time, even when there is complete Scopus coverage. Citations per Publication calculates an average value, and these types of calculations are strongly influenced by outlying publications in a small data set.

- There is a concern that excessive self-citations may be artificially inflating Citations per Publication. Users can judge whether the level of self-citations is higher than expected by deselecting the "Include self-citations" option.

- Uncovering the performance of publications in the very early stages of a new strategy, or of early- career researchers, where the short time that has passed since publication will reduce the reliability of basing decisions on citation information. It is advised to use metrics like Scopus Source Title Count or Collaboration to account for this.

- The person who will use the data does not like to see a line that "dips" in recent years. This typically happens with Citations per Publication because recent publications have had little time to receive citations. Users are advised to use Field-Weighted Citation Impact to avoid this drop, if it is of concern.

Useful partner metrics are:

- **Field-Weighted Citation Impact** is a logical complement to Citations per Publication, considering behavioral differences between disciplines

- **Field-Weighted Citation Impact**, **Outputs in Top Citation Percentiles**, and **Publications in Top Journal Percentiles** avoid display of the "dip" in recent years due to little time having passed to receive citations since publication

- **Cited Publications**, which provides a measure of the reliability that an entity's publications will subsequently be used to support other research, but that is not affected by a high number of citations received by 1 or a few publications

- **Collaboration Impact** and **Academic-Corporate Collaboration Impact** also measure Citations per Publication, and are a complement. They focus on subsets of publications within a data set with collaboration characteristics.

- **The set of "time-independent metrics"** that provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Subject Area Count, Scopus Source Title Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

ELSEVIER

5.5.6 Metric: Cited Publications

Cited Publications in SciVal indicates the citability of a set of publications: how many of this entity's publications have received at least 1 citation?

Cited Publications is a:

- Citation Impact metric

- "Power Metric" when the "Total value" option is selected, but not when the "Percentage" option is selected

SciVal often displays Cited Publications in a chart or table with years. These years are always the years in which items were published, and do not refer to the years in which citations were received.

This metric is useful to:

- Benchmark the extent to which publications are built on by subsequent work

- Compare the influence of publications of entities of different sizes, but in related disciplines, such as large and small countries, or a research team with individual Researchers within that team

  – It is advised to select the "Percentage" option when comparing entities of different sizes, to normalize for this variable.

- Demonstrate the excellence of entities that produce consistent work that is reliably cited, regardless of the number of citations received

- Account for the positive impact of a few very highly cited papers on apparent performance, which affect Citation Count and Citations per Publication

This metric should be used with care when:

- Benchmarking the extent to which the publications of entities with distinct disciplinary profiles are built on by subsequent work:

  – Teams working in parasitology, for instance, may experience a relatively short delay between publishing and the receipt of citations because of the high frequency of publishing and citing, relative to teams working in mathematical modeling where publication behavior may lead to a relatively long delay between publishing and the receipt of citations.

  – It is not advisable to use this metric to compare entities in distinct disciplines without taking these differences into account.

  – When comparing entities made up of a mixture of disciplines, such as an Institution or a Group of Institutions, it is advised to apply the Research Area filter to focus on one field that is common between all the entities.

- Understanding the magnitude of the number of citations received by an entity's publications. Users are advised to use Citation Count or Citations per Publication to communicate this information

- Entities are small and there may be gaps in their output within the Scopus coverage:

  – A single missing publication from a small data set may have a significant negative impact on apparent visibility. For example, an academic's performance may suffer due to gaps in items they have published, as well as gaps in publications citing the publications that are indexed.

  – The only way to account for this is to be vigilant. Consider also limiting the use of Cited Publications to comparing larger data sets in the same discipline in which potential gaps in the database coverage likely have a similar effect on all entities being viewed and do not invalidate the comparison.

- There is a concern that excessive self-citations may be artificially inflating the proportion of cited publications. Users can judge whether the level of self-citations is higher than expected by deselecting the "Include self-citations" option.

- Uncovering the performance of publications in the very early stages of a new strategy, or of early-career researchers, where the short time that has passed since publication will reduce the reliability of based on decisions on citation information

Useful partner metrics are:

- **Citations Count** and **Citations per Publication**, which indicate the magnitude of citation impact, and which will be positively impacted by one or a few publications that have received a very high number of citations

- **The set of all other "Power Metrics"** whose value tends to increase as the entity becomes bigger: Scholarly Output, Subject Area Count, Citation Count, Number of Citing Countries, Collaboration ("Total value"), Academic-Corporate Collaboration ("Total value"), Outputs in Top Citation Percentiles ("Total value"), Publications in Top Journal Percentiles ("Total value"), and *h*-indices

- **The set of "time-independent metrics"** that provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Subject Area Count, Scopus Source Title Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

5.5.7 Metric: *h*-indices ❄

Please see for documentation regarding *h*-indices.

5.5.8 Metric: Number of Citing Countries

Number of Citing Countries in SciVal indicates the geographical visibility of an entity's publications: from how many distinct countries have this entity's publications received citations?

Number of Citing Countries is a:

- Citation Impact metric

- "Power Metric"

SciVal often displays Number of Citing Countries in a chart or table with years. These years are always the years in which items

ELSEVIER

were published, and do not refer to the years in which citations were received.

This metric is useful to:

- Compare smaller entities, such as Groups of Researchers and Publication Sets, where differences in the number of citing countries are most likely to be evident

- Provide impressive figures to showcase the performance of a relatively large entity, when this metric is likely to give high numbers

- Benchmark the geographical visibility of the publication portfolios of related entities, such as:

  - Collaboration networks in a given field of research

  - Researchers in a common field of research and with similar career lengths

  - Scenario models of a research institute, created to investigate the effect of recruiting different researchers

- Provide evidence of extensive geographical appeal of a body of work by indicating the diversity of the geographical sources of citations

- Look at publishing activity in a way that is difficult to manipulate

This metric should be used with care when:

- Comparing large entities, such as Institutions and Groups of Countries, which will likely publish so many publications that receive so many citations, that the number of citing countries will approach the maximum and differences may not be visible. Users are advised to narrow their focus to a slice of a large entity when using this metric, such as by applying the Research Area filter.

- Benchmarking entities of obviously different sizes, such as institutions and small departments, when this "Power Metric" will most closely reflect entity size rather than differences in global visibility. It is not advised to use this metric to compare entities of different sizes.

- Benchmarking the collaboration of entities in different disciplines:

  - The international appeal of publications from different disciplines may be inherently different, such as between national literature and chemistry, or local history and computer science.

  - It is not advisable to use this metric to compare entities in distinct disciplines without accounting for these differences.

  - When comparing entities made up of a mixture of disciplines, such as an Institution or a Country, it is advised to apply the Research Area filter to focus on one field that is common between all the entities.

- Understanding the magnitude of the number of citations received by an entity's publications. Users are advised to use Citation Count or Citations per Publication to communicate information about magnitude.

- Entities are small and there may be gaps in their output within the Scopus coverage:

  - A single missing publication from a small data set may have a significant negative impact on apparent global visibility, whereas the effect of 1 or a few missing publication(s) from a large data set may be acceptable.

  - The only way to account for this is to be vigilant, particularly when looking at small data sets such as an early-career researcher. Consider also limiting the use of Number of Citing Countries to comparing somewhat larger data sets in the same discipline in which potential gaps in the database coverage likely have a similar effect on all entities being viewed and do not invalidate the comparison.

- Investigating the performance of publications in the very early stages of a new strategy, or of early-career researchers, where the short time that has passed since publication will reduce the reliability of basing decisions on citation information. It is advised to use Subject Area Count or other time-independent metrics in these situations.

Useful partner metrics are:

- **Citation Count** and **Citations per Publication**, which communicate information about the magnitude of the number of citations received

- **Collaboration**, which shows the extent of international co-authorship of an entity's scholarly output and is a complement to Number of Citing Countries' view on geographical visibility

- **The set of all other "Power Metrics"** whose value tends to increase as the entity becomes bigger: Scholarly Output, Subject Area Count, Scopus Source Title Count, Citation Count, Cited Publications ("Total value"), Collaboration ("Total value"), Academic-Corporate Collaboration ("Total value"), Outputs in Top Citation Percentiles ("Total value"), Publications in Top Journal Percentiles ("Total value"), and h-indices

- **The set of "time-independent metrics"** that provide useful, reliable information immediately upon publication and do not rely on the passing of time for useful data to accumulate: Scholarly Output, Subject Area Count, Scopus Source Title Count, Collaboration, Academic-Corporate Collaboration, and Publications in Top Journal Percentiles

5.5.9 Metric: Collaboration Impact ❄

See Metric: Collaboration Impact on page 34

5.5.10 Metric: Academic-Corporate Collaboration Impact ❄

See Metric: Academic-Corporate Collaboration on page 36

5.5.11 Metric: Citing-Patents Count

See Metric: Citing Patents Count on page 54

5.5.12 Metric: Patent-Cited Scholarly Output

See Metric: Patent-Citations Count on

5.5.13 Metric: Patent-Citations Count

See Metric: Patent-Citations Count on

5.5.14 Metric: Patent-Citations per Scholarly Output

See Metric: Patent-Citations per Scholarly Output on

5.6 Economic Impact

SciVal identifies and counts citations which research papers have received from patents. From the perspective of a research publication, these would be "forward citations", indicating whether the research results have subsequently been used in the patent world. It is important to remember that patents are published and can only become available for use in research metrics around eighteen months after the application date.

- Citation of scholarly output in patents indicates a connection between academia and industry.

- SciVal's patent-related metrics serve as a tool to help detect and demonstrate research impact.

- SciVal looks at the citations of scholarly output in patents and links to both the citing patents and cited articles.

5.6.1 Metric: Academic-Corporate Collaboration ❄

See Metric: Academic-Corporate Collaboration on

5.6.2 Metric: Academic-Corporate Collaboration Impact ❄

See Metric: Academic-Corporate Collaboration Impact on

5.6.3 Metric: Citing-Patents Count

This is the count of patents citing the scholarly output published by the entity (e.g. a university) in which you are looking. The count of patents may be higher than the number of scholarly outputs cited, since multiple patents could refer to the same piece of output. The count of outputs may be higher than the number of patents since one patent can refer to multiple scholarly outputs.

Example: 200 patents have cited articles published by Athena University over the past five years. The Citing-Patent Count metric equals 200.

This metric is useful to:

- Gain an understanding of the impact of research in the creation of products by seeing the total number of patents received by the research

- Provide information about the economic impact of research

This metric should be used with care when:

- Comparing disciplines that are not likely to have research referenced in a patent such as Arts and Humanities

Evaluating a country that is not covered by the patent offices in SciVal

Useful partner metrics are:

- **Patent-Cited Scholarly Output**

- **Patent-Citations Count**

- **Patent-Citations per Scholarly Output**

5.6.4 Metric: Patent-Cited Scholarly Output

This is the count of scholarly output published by an entity (e.g. a university) that have been cited in patents.

Example: 400 publications from Athena University have been cited by patents. The Patent-Cited Scholarly Output metrics equals 400.

This metric is useful to:

- Gain an understanding of how much research is being used in the creation of products by seeing the total number of outputs that received patents

- Provide information about the economic impact of research

This metric should be used with care when:

- Comparing disciplines who are not likely to have research referenced in a patent such as Arts and Humanities

- Evaluating a country that is not covered by the patent offices included in SciVal

Useful partner metrics are:

- **Patent-Cited Scholarly Output**

- **Patent-Citations Count**

- **Patent-Citations per Scholarly Output**

5.6.5 Metric: Patent-Citations Count

This is the total count of patent citations received by the entity (e.g. a university).

Example: Athena University has been cited 600 times by patents over the past five years. From our example this means that the 400 publications from Athena University's 400 publications have been cited 600 times by the 200 patents.

This metric is useful to:

- Gain an understanding of how much research is being used in the creation of products by seeing the total times the outputs have received patent citations

- Provide information about the economic impact of research

This metric should be used with care when:

- Comparing disciplines who are not likely to have research referenced

in a patent such as Arts and Humanities

- Evaluating a country that is not covered by the patent offices used in SciVal

Useful partner metrics are:

- **Patent-Cited Scholarly Output**
- **Patent-Citations Count**
- **Patent-Citations per Scholarly Output**

5.6.6 Metric: Patent-Citations per Scholarly Output

This is the average patent-citations received per 1,000 scholarly outputs published by the entity (e.g. a university). i.e. the patent-citation counts divided by the total scholarly output of the university for that period and multiplied by 1,000.

Example: If Athena University published 10,000 publications in a five-year period, their patent-citations per scholarly output would be (600/10,000) x 1,000 = 60. We look at this metric per 1,000 publications because otherwise the typical average patent-citations per output is a small number that is harder to interpret.

This metric is useful to:

- Gain an understanding of how much, on average, research is being used in the creation of products
- Provide information about the economic impact of research

This metric should be used with care when:

- Comparing disciplines who are not likely to have research referenced in a patent such as Arts and Humanities
- Evaluating a country that is not covered by the patent offices in SciVal

Useful partner metrics are:

- **Patent-Cited Scholarly Output**
- **Patent-Citations Count**
- **Patent-Citations per Scholarly Output**

5.7 Societal Impact

5.7.1 Metric: Mass Media ❄

Mass Media refers to the total number of times that the media referred to researchers of the selected institution(s). Mass Media considers media articles from 2011 onwards and is currently focused on media articles in the English language only. Mass Media mentions have a different range of coverage depending on the medium: print media sources are covered from 2011 onwards whereas online media sources are covered from 2014 onwards. The entire database of Mass Media content consists of[18]:

- LexisNexis print archive (Terms and Conditions

http://www.lexisnexis.co.uk/en-uk/terms.page): 2011 till October 2015

- LexisNexis Metabase print news (Terms and Conditions http://www.lexisnexis.co.uk/en-uk/terms.page): since November 2015
- LexisNexis Metabase online news: since September 2013

Several media types are covered, including:

- Online News: a Web based news publication
- Print: clippings or text that was originally made available in print
- Blog: an interactive website made up of entries or posts displayed in reverse chronological order, as well as reader comments on the posts.
- Comment: comments which are posted on blogs or some news articles.

Media mentions are delivered via a near real-time feed from the LexisNexis Metabase portal to NewsFlo[19]. Due to the large search query of finding all Scopus author-affiliation combinations, Mass Media in SciVal is updated every two months.

To allocate articles at an institution level, all news articles in the database are clustered by text-similarity and then matched against the Scopus database. If a Scopus author name and one of his/her affiliations (current/previous) are both found in at least one of the articles of the cluster, then the cluster is validated and all news articles will be assigned to this Scopus author ID and Scopus affiliation ID. The Scopus affiliation ID values are aggregated to calculate media mentions at the institutional level.

Mass Media can be filtered by the media source category, the options being:

- Academic: News items from educational institutions, e.g. schools and universities
- Corporate: Corporate website press release pages, e.g. McDonalds, Shell
- General: News items from sources that are yet to be assigned a source category
- Government: News and information from governments and government departments
- Journal: Periodical publications, typically focused on science, technology or professions
- Local: News from local and regional news sources, e.g. The Alaska Star, Bath Chronicle
- Miscellaneous: News from sources that do not fit into any of the other source categories
- Organization: News from organizations such as charities, political parties, NGOs

18 https://www.elsevier.com/solutions/newsflo/sources
19 https://www.elsevier.com/solutions/newsflo

- Press Wire: Designated press release and press wire sources, e.g. Business Wire

- Trade: News items from designated industry, profession or technology focused sources, e.g. Financial Review, McKinsey Quarterly, Oil and Gas Journal

In addition, filtering of Mass Media is possible based on the assignment to one of five tiers indicating the source exposure.

These source tiers are:

- Tier 1: Internationally recognized
- Tier 2: Regionally recognized
- Tier 3: Nationally recognized
- Tier 4: Locally recognized
- Tier 5: Local interest

The tiers are assigned by LexisNexis. Several aspects are considered when assigning a source tier, the most important of which are: coverage around the world (how many regions a medium is published across), importance (leading news source in its area) and the type of news. These factors may not necessarily carry the same weight for every source.

For example, a very renowned source that only publishes country wide could be classified with a source tier of 1, because it is so well known in its country of publication. Whereas a source that published globally may be given a source tier of 4 because it is not very well known, and only publishes industry-specific content for example.

Mass Media is a:

- Snowball metric
- Power Metric

This metric is useful to:

- Showcase the engagement of the broader public, and not only of the research community

- Indicate some degree of societal impact

- Gain an understanding which research outputs grab the media's attention

- Discover media outlets that are discussing an institution's outputs

- Understand the level of global exposure mentioned outputs received

- Give an early indication of the interest in output that has recently become available

This metric should be used with care when:

- Comparing across disciplines. Some disciplines naturally capture the public's attention more than others, but this does not mean that they have less social significance.

- Comparing across years. The database of Mass Media content is subject to fluctuations over time, with license agreements being

time-limited, and media sources being added or removed from the database. In addition, the discovery of Mass Media mentions is dependent on the journalism culture itself. For example, it may be that writing styles are adapting to no longer mention author names and/or affiliations, which are the basis for discovering Mass Media mentions. Often, the media do not mention the output they are referring to in a way that can be automatically recognized, so that the database coverage will be less than 100% percent of an institution's productivity. It is also possible that the media are using name variants of institutions that don't match against the database, or that the media report less/more on science altogether.

- Comparing across regions. As the database currently only considers media output in the English language, mentions in local languages will not be detected and counted towards Mass Media. This might skew the data towards regions in which English is the native language.

Useful partner metrics are:

- **Media Exposure**, which indicates the number of mentions weighted by type of publication, demographics and audience reach

- **Field-Weighted Mass Media**, which helps take away the effect different disciplines have on media exposure

5.7.2 Metric: Media Exposure

- Media Exposure indicates the number of media mentions weighted by type of publication, demographics and audience reach. (See Metric: Mass Media on page 83 for more information on how media mentions are assigned). The weighting is assigned based on the source tier:

  - Internationally recognized = total count of Mass Media in that tier **x1**

  - Regionally recognized = total count of Mass Media in that tier **x0.5**

  - Nationally recognized = total count of Mass Media in that tier **x0.3**

  - Locally recognized = total count of Mass Media in that tier **x0.2**

  - Local interest = total count of Mass Media in that tier **x0.1**

This metric is useful to:

- Gain an understanding of the relative impact an institution's research outputs have on the media

- Understand to what extend mentions were found in sources that are internationally recognized, regionally recognized, nationally recognized, locally recognized or of local interest

- Showcase the engagement of the broader public, and not only of the research community

This metric should be used with care when:

- Comparing across disciplines. Some disciplines naturally capture the public's attention more than others, but this does not mean that they have less social significance.

- Comparing across years. The database of Mass Media content is subject to fluctuations over time, with license agreements being time-limited, and media sources being added or removed from the database. In addition, the discovery of Mass Media mentions is dependent on the journalism culture itself. For example, it may be that writing styles are adapting to no longer mention author names and/or affiliations, which are the basis for discovering Mass Media mentions. Often, the media do not mention the output they are referring to in a way that can be automatically recognized, so that the database coverage will be less than 100% percent of an institution's productivity. It is also possible that the media are using name variants of institutions that don't match against the database, or that the media report less/more on science altogether.

- Comparing across regions. As the database currently only considers media output in the English language, mentions in local languages will not be detected and counted towards Mass Media. This might skew the data towards regions in which English is the native language.

Useful partner metrics are:

- **Mass Media**, which indicates the total number of media mentions received

- **Field-Weighted Mass Media**, which helps take away the effect different disciplines have on media exposure

5.7.3 Metric: Field-Weighted Mass Media

Field-Weighted Mass Media is the ratio of Mass Media mentions relative to the expected world average for the subject field, publication type and publication year. (See Metric: Mass Media on page 83 for more information on how media mentions are assigned).

Similar publications are those publications in the database that have the same publication year, publication type, and discipline, as represented by the Scopus classification system.

- Field-Weighted Mass Media of 1.00 indicates that the entity's publications have been mentioned in the media exactly as would be expected based on the global average for similar publications; the Field-Weighted Mass Media of "World", or the entire Scopus database, is 1.00.

- Field-Weighted Mass Media of more than 1.00 indicates that the entity's publications have been mentioned in the media more than would be expected based on the global average for similar publications; for example, 2.11 means 111% more than the world average.

- Field-Weighted Mass Media of less than 1.00 indicates that the entity's publications have been mentioned in the media less than would be expected based on the global average for similar publications; for example, 0.87 means 13% less than the world average.

This metric is useful to:

- Benchmark entities regardless of differences in their size, disciplinary profile, age, and publication-type composition

- Easily understand the prestige of an entity's media exposure by observing the extent to which its Field-Weighted Mass Media is above or below the world average of 1.00

- Showcase the engagement of the broader public, and not only of the research community.

This metric should be used with care when:

- Information about the magnitude of the number of media mentions received by an entity's publications is important. In these situations, it is advised to use Mass Media or Media Exposure.

- Demonstrating excellent media reception to those who prefer to see high numbers; Mass Media or Media Exposure would be more suitable in these circumstances

- Comparing across regions. As the database currently only considers media output in the English language, mentions in local languages will not be detected and counted towards Mass Media. This might skew the data towards regions in which English is the native language.

Useful partner metrics are:

- **Mass Media**, which indicates the total number of media mentions received

- **Media Exposure**, which indicates the number of mentions weighted by type of publication, demographics and audience reach

5.8 Examples

5.8.1 Example 2: Scholarly Output, Subject Area Count and Scopus Source Title Count

Scenario: The user would like to calculate the Scholarly Output, Journal Count, or Scopus Source Title Count
of an entity that consists of 6 publications, and has selected the following viewing and calculation options.

Selected Publication Year Range	2008 to 2012
Selected Publication Types	Articles, reviews and conference papers
Selected Research Area	Agricultrual and Biological Sciences

	Entity with 6 Publications						
Publication Identity	Publication 1	Publication 2		Publication 3		Publication 4	Publication 5	Publication 6
Publication year	2007	2009		2009		2008	2010	2010
Journal in which publication is published	*Bioscience Journal*	*Biology and Environment*		*Biology and Environment*		*Biology and Environment*	*Archiv für Lebensmittelhygiene*	*Archiv für Lebensmittelhygiene*
Publication type	Article	Article		Article In Press		Review	Article	Report
Journal sub-category(s)	General Agricultural and Biological Sciences	Animal Science and Zoology	Aquatic Science	Insect Science	Environmental Chemistry	Agronomy and Crop Science	Horticulture	Organic Chemsitry
Journal main-category(s)	Agricultural and Biological Sciences	Agricultural and Biological Sciences	Agricultural and Biological Sciences	Agricultural and Biological Sciences	Environmental Science	Agricultural and Biological Sciences	Agricultural and Biological Sciences	Chemistry
Do publications match user-selected options?								
Step 1 Does the Scopus Source Title Count of the publication match the selected research area?	Yes	Yes	Yes	Yes	No	Yes	Yes	No
Step 2 Does the publication fall in the selected publication year range?	No	Yes		Yes		Yes	Yes	Yes
Step 3 Does the publication type match the selected publication type?	Yes	Yes		No		Yes	Yes	No
Step 4 Does the publication pass each of steps 1, 2 and 3?	No	Yes		No		Yes	Yes	No

Question: How do I calculate Scholarly Output?	**Answer:** Count the number of publications that receives a "yes" in Step 4.	**Scholarly Output = 3**
Question: How do I calculate Scopus Source Title Count?	**Answer:** Note unique Scopus Source Title Count for those publications that received a "yes" in step 4. • Count the number unique Journal Titles.	**Journal Count = 2** Remove Duplicates Journal Titles of those publications that pass the filters. *Biology and Environment, Archiv für Lebensmittelhygiene*
Question: How do I calculate Scopus Source Title Count (Main Categories)?	**Answer:** Look up Main Categories. • Remove duplicates. • Count the number of unique Main Categories	**Journal Main Category Count = 1** Remove duplicate Main Category name from the publications that passed the filters. *Agricultural and Biological Sciences*
Question: How do I calculate Scopus Source Title Count (Sub-categories)	**Answer:** Look up Sub-Categories. • Remove duplicates. • Count the number of unique sub-categories.	**Journal Sub-Category Count = 4** Remove duplicate Sub-Category names from those publications that passed the filters. *Agronomy and Crop Science, Animal Science and Zoology, Aquatic Science, Horticulture*

5.8.2 Example 3: Citation Count, Cited Publications and Citations per Publication

Scenario: The user would like to calculate the Citation Count, Cited Publications, or Citations per Publication of an entity that consists of 6 publications, and has selected the following viewing and calculation options.

Selected Publication Year Range	2005 to 2013
Selected Publication Types	Articles, reviews and editorials
Selected Research Area	Medicine

Publication Identity	Publication 1	Publication 2		Publication 3		Publication 4	Publication 5	Publication 6
Publication year	2008	2007		2005		2004	2008	2010
Publication type	Review	Editorial		Article		Article	Article	Review
Total citations received by this publication	0	4		7		9	0	4
Journal sub-category(s)	Anthropology	Emergency Medicine	Management Science and Operations Research	Anatomy	Information Systems and Management	Immunology and Allergy	General Medicine	Emergency Medicine
Journal main-category(s)	Social Sciences	Medicine	Decision Sciences	Medicine	Decision Sciences	Medicine	Medicine	Medicine
Do publications match user-selected options?								
Step 1	Does the Scopus Source Title Count of the publication match the selected research area?	No	Yes	No	Yes	No	Yes	Yes	Yes
Step 2	Does the publication fall in the selected publication year range?	Yes	Yes		Yes		No	Yes	Yes
Step 3	Does the publication type match the selected publication type?	Yes	Yes		Yes		Yes	Yes	No
Step 4	Does the publication pass each of steps 1, 2 and 3?	No	Yes		Yes		Yes	Yes	No

		Publication 1	Publication 2	Publication 3	Publication 4	Publication 5	Publication 6
Question: How do I calculate Scholarly Outputs?	**Answer:** Count the number of publications that receives a "yes" in Step 4.	**Scholarly Output = 4**					
Question: How do I calculate Citation Count?	**Answer:** Retrieve total citations received for the publications that received a "yes" in step 4.	N/A	4	7	N/A	0	4
	Sum the citations received by those publications that received a "Yes" in step 4.	**Citation Count = 15**					
Question: How do I calculate Cited Publications?	**Answer:** Have the publications that passed step 4 received at least 1 citation.	N/A	Yes	Yes	N/A	No	Yes
	For "Absolute number" display option, sum the publications that have received at least 1 citation.	**Cite Publications = 3**					
	For "Percentage" display option, divide "Absolute number" by Scholarly Output.	**Cited Publications = 75%**					
Question: How do I calculate Citations per Publication?	**Answer:** Divide "Citation Count" by "Scholarly Output"	**Citations per Publications = 3.8**					

5.8.3 Example 4: Number of Citing Countries

Scenario: The user would like to calculate the Number of Citing Countries of an entity that consists of 6 publications. They have not selected any viewing or calculation options. Say that this entity has received 6 citations from publications A, B, C, D, E and F.

	Entity with 6 Publications					
	Publication 1	Publication 2	Publication 3	Publication 4	Publication 5	Publication 1
Cited by: Publication A	Yes				Yes	
Cited by: Publication B			Yes		Yes	
Cited by: Publication C	Yes	Yes			Yes	Yes
Cited by: Publication D	Yes	Yes		Yes		Yes
Cited by: Publication E	Yes	Yes		Yes		Yes
Cited by: Publication F		Yes			Yes	

Scenario: The citing publications A, B, C, D, E and F have the following affiliation information:

Citing Publication	Authors	Institutions	Country
Publication A	A_2	I_4	C_2
Publication B	A_1 A_3 A_3	I_1 I_4 I_2	C_1 C_2 C_1
Publication C	A_1 A_1 A_2	I_1 I_3 I_2	C_1 C_2 C_1
Publication D	A_1 A_4	I_1 I_1	C_1 C_1
Publication E	A_1 A_3 A_5	I_1 I_5 I_1	C_1 C_3 C_1
Publication F	A_1 A_3 A_2	I_1 I_8 I_1	C_1 C_4 C_1

Question: How do I calculate the number of Citing Countries?

Answer: Count the number of distinct countries in the affiliations of the citing publications.

Number of Citing Countries = 4

5.8.4 Example 5: Field-Weighted Citation Impact

Scenario: The user would like to calculate the Field-Weighted Citation Impact of an entity that consists of 3 publications. They have not selected any viewing or calculation options.

		Entity with 3 Publications			
	Publication Identity	Publication 1	Publication 2		Publication 3
	Publication year (pub year)	2009	2010		2013
	Publication type	Article	Review		Erratum
	Journal category(s)	Immunology	Immunology	Parasitology	Parasitology
Step 1	• Compute number of citations received by publications in entity.	2	12		0
	• Actual citations received in pub year	3	23		N/A (Example prepared in 2013)
	• Actual citations received in 1st year after pub year	13	28		N/A (Example prepared in 2013)
	• Actual citations received in 2nd year after pub year	23	34		N/A (Example prepared in 2013)
	• Actual citations received in 3rd year after pub year	2+3+13+23	12+23+28+45		
	• Actual citations received by the individual publication in pub year plus following 3 year	41	108		0
Step 2	• Compute expected number of citations received by similar publications.	7,829.6	1,349.8	161.9	8.3
	• Number of publications in database published in same year, of same type, and within same discipline as publication 1, 2, or 3	141,665.2	35,770.8	2,161.5	0.0
	• Total citations received in pub year plus 3 years by all publications in database published in same year, of same type, and within same journal category(ies) as Publication 1, 2, or 3	141,665.2 / 7,829.6	35,770.8 / 1,349.8	2,161.5 / 161.9	0.0 / 8.3
	• Average citation per publication for all publications in database published in same year, of same type, and within same subject category as Publication 1, 2, or 3	18.09	26.50	13.35	0.00
Step 3	• Use harmonic mean to compute expected number of citations for publications covered in multiple categories.	2 / (1/26.5 + 1/13.4)			
	• Combined average citations per publication for publications indexed in multiple journal categories	17.76			
Step 4	Compute ratio of actual (result of step 1) to expected (result of step 2 or 3) citations for each of Publications 1, 2 and 3.	41 / 18.1 2.27	108 / 17.76 6.08		No citations received or expected => 0 0.00
Step 5	Take arithmetic mean of the results of step 4 to calculate Field-Weighted Citation Impact for this entity	Arithmetic Mean (2.27 + 6.08 + 0.00) / 3 **Field-Weighted Citation Impact = 2.78**			

5.8.5 Example 6: Collaboration, Collaboration Impact, Academic-Corporate Collaboration, and Academic-Corporation Collaboration Impact

Scenario: The user would like to calculate the Collaboration, Collaboration Impact, Academic-Corporate Collaboration, or Academic-Corporate Collaboration Impact of an entity that consists of 6 publications.

Selected Publication Year Range	2005 to 2012
Selected Publication Types	Articles, reviews and editorials
Selected Research Area	Medicine

		Entity with 6 Publications				
	Publication Identity	Publication 1		Publication 2		
	Publication year	2011		2009		
	Total citations received by this publication	8		11		
	Publication type	Letter		Review		
	Journal sub-category(s)	Anatomy		Dermatology		
	Journal main-category(s)	Medicine		Medicine		
	Authors	Author 1	Auther 2	Author 1	Author 3	
	Institutions	Institution 1	Institution 1	Institution 1	Institution 4	Institution 2
	Countries	Country 1	Country 2	Country 1	Country 2	Country 1
	Is the institution academic?	Yes	Yes	Yes	No	No
	Is the institution corporate?	No	No	No	No	Yes
Step 1	Does the Scopus Source Title Count of the publication match the selected research area?	Yes		Yes		
Step 2	Does the publication fall in the selected publication year range?	Yes		Yes		
Step 3	Does the publication type match the selected publication type?	No		Yes		
Step 4	Does the publication pass each of steps 1, 2 and 3?	No		Yes		

Question: How do I calculate Scholarly Output?	**Answer:** Count the number of publications that received a "yes" in step 4.	
Question: How do I calculate Collaboration?	**Answer:** Assign collaboration type to each publication that received a "yes" in step 4 using the decision tree. • Does the publication have a single author? If yes, assign as single authorship. • Do the unassigned publications have more than one country in the affiliation information? • Do the unassigned publications have more than one institution in the affiliation information? If yes, assign as National Collaboration. • Assign all remaining publications as Institutional Collaboration.	International Collaboration
	For "Absolute number" display option, sum the publications within each group.	Single Authorship – 1 International Collaboration – 2 National Collaboration – 0 Institutional Collaboration – 1
	For "Percentage" display option, divide "Absolute Number" within each group by Scholarly Output of the Entity.	Single Authorship – (1/4)*100 – 25% International Collaboration – (2/4)*100 – 50% National Collaboration Impact – 0/0 – 0 Institutional Collaboration Impact – 4/1 – 4
Question: How do I calculate Collaboration Impact?	**Answer:** Divide total citations received by the publications by absolute number of each type of collaboration.	Single Authorship Impact – 1/4 – 1.0 International Collaboration Impact – (11+6)/2 – 8.5 National Collaboration Impact – 0/0 – 0 Institutional Collaboration Impact – 4/1 – 4
Question: How do I calculate Academic-Corporate Collaboration?	**Answer:** Do the publications that received a "yes" in step 4 have both an academic and a corporate institution in the affiliation information?	Yes
	For "Absolute number" display option, sum the publications with and without academic-corporate collaboration.	With Collaboration – 1 Without Collaboration – 3
	For "Percentage" display option, divide "Absolute Number" within each group by Scholarly Output of the entity.	With Collaboration – (1/4)*100 – 25% Without Collaboration – (3/4)*100 – 75%
Question: How do I calculate Academic-Corporate Collaboration Impact?	**Answer:** Divide total citations received by the publications by absolute number of each type if collaboration.	Impact with collaboration – 11/1 – 11 Impact without collaboration – (1+4+6) / 3 – 3.7

Publication 3			Publication 4		Publication 5			Publication 6		
2011			2011		2010			2003		
1			4		6			12		
Article			Editorial		Review			Article		
Anatomy			Internal Medicine		General Medicine			Emergency Medicine		
Medicine			Medicine		Medicine			Medicine		
Author 1			Author 1	Author 4	Author 1		Author 4	Author 1		Author 2
Institution 1	Institution 3	Institution 5	Institution 1	Institution 1	Institution 1	Institution 1	Institution 1	Institution 1	Institution 2	Institution 1
Country 1	Country 2	Country 3	Country 1	Country 1	Country 1	Country 3	Country 1	Country 1	Country 1	Country 1
Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
No	No	No	No	No	No	No	No	No	Yes	No
Yes			Yes		Yes			Yes		
Yes			Yes		Yes			No		
No			Yes		Yes			Yes		
No			Yes		Yes			No		

Single Authorship	Institutional Collaboration	Institutional Collaboration	
No	No	No	

5.8.6 Example 7: Outputs in Top Citation Percentiles

Scenario: The user would like to calculate the Outputs in the Top Citation Percentiles of an entity that consists of 6 publications, and has selected the following viewing and calculation options:

Selected Publication Year Range	2004 to 2013
Selected Publication Types	Articles, reviews and editorials
Selected Research Area	Chemistry
Selected Percentile Level	10%
Selected Data Universe	World

Entity with 6 Publications

		Publication 1	Publication 2
	Publication Identity	Publication 1	Publication 2
	Publication year	2011	2009
	Total citations received by this publication	10	42
	Publication type	Article	Review
	Journal sub-category(s)	Organic Chemistry	Pharmaceutical Science
	Journal main-category(s)	Chemistry	Pharmacology, Toxicology and Pharmaceuticals
Do publications match user-selected options?			
Step 1	Does the journal category of the publication match the selected research area?	Yes	No
Step 2	Does the publication fall in the selected publication year range?	Yes	Yes
Step 3	Does the publication type match the selected publication type?	Yes	Yes
Step 4	Does the publication pass each of steps 1, 2 and 3?	Yes	No

Question: How do I calculate Scholarly Output?	**Answer:** Count the number of publications that received a "yes" in step 4.	**Scholarly Output – 4**	
Question: How do I calculate Outputs in Top Citation Percentiles?	**Answer:** Look up the 10% citation threshold for this Publication Year of Publications that received a "yes" in step 4.	8	N/A
	Was the publication cited at least as many times as the threshold?	Yes	N/A
	For "Absolute number" display option, count the publications that received a "yes" in the previous step.	**Output in Top Citation Precentiles = 3**	
	For "Percentage" display option, divide "Absolute Number" by the Scholarly Output of the entity.	**Outputs in Top Citation Precentiles = (3/4)*100 = 75%**	

	Publication 3	Publication 4	Publication 5	Publication 6
	2011	2011	2010	2003
	15	4	6	12
	Editorial	Editorial	Review	Article
	Geochemistry and Petrology	Internal Medicine	General Medicine	Emergency Medicine
	Earth and Planetary Sciences	Medicine	Medicine	Medicine
	No	Yes	Yes	Yes
	Yes	Yes	Yes	No
	Yes	Yes	Yes	Yes
	No	Yes	Yes	No

	Publication 3	Publication 4	Publication 5	Publication 6
	N/A	39	13	39
	N/A	No	Yes	Yes

This table shows the number of times a publication must be cited to be in the top 10% for its Publication Year, based on arbitrary thresholds.

10% World Citation Thresholds

Year	Citations	Year	Citations	Year	Citations
1996	44	2003	42	2010	13
1997	45	2004	39	2011	8
1998	47	2005	33	2012	4
1999	49	2006	30	2013	1
2000	49	2007	26	2014	0
2001	45	2008	22		
2002	44	2009	18		

5.8.7 Example 8: Publications in Top Journal Percentiles

Scenario: The user would like to calculate the Publications in Top Journal Percentiles of an entity that consists of 6 publications, and has selected the following viewing and calculation options.

Selected Journal Metric	SNIP
Selected Percentile Level	10%

		Entity with 6 Publications					
	Publication Identity	Publication 1	Publication 2	Publication 3	Publication 4	Publication 5	Publication 6
	Publication year	2005	2006	2007	2007	2006	2006
	Journal in which publication is published	*Preception*	*Journal of the American Medical Association*	*Vision Research*	*Nature*	*Vision Research*	*Biophysics*
Step 1	Retrieve journal's SNIP value for the Publication Year	0.851	8.930	1.130	7.100	1.292	0.036
Step 2	Look up the 10% SNIP threshold for the Publication Year	1.590	1.573	1.566	1.566	1.573	1.573
Step 3	Is this Journal's SNIP (step 1) at leas as large as the 10% SNIP threshold (step 2)?	No	Yes	No	Yes	No	No

Question: How do I calculate Scholarly Output?	Answer: Count the number of publications in the entity.	Scholarly Output = 6
Question: How do I calculate Publications in Top Journal Percentiles?	Answer: For "Total value" display option, count the publications that are published in the top 10% of journals.	Publications in Top 10% Journals Percentiles = 2
	For "Percentage" display option, divide "Total value" by the Scholarly Output of the entity.	Publications in Top 10% Journals Percentiles = (2/6)*100 = 33.3%

SNIP Percentile Thresholds

Percentile Level	2005	2006	2007
1%	4.116	4.024	4.036
5%	2.089	2.047	2.034
10%	1.590	1.573	1.566
25%	1.072	1.061	1.045

5.8.8 Example 9: *h*-indices

Scenario: The user would like to calculate the h-indices of an entity that consists of 10 publications. They have not selected any viewing or calculation options.

		Entity with 10 Publications									
	Publication year	2005	2003	1998	1997	1997	1998	2008	1997	2001	2002
	Total citations received by this publication	2	5	4	12	4	12	4	6	2	5
Order and rank the publications											
Step 1	Sort the publications by their citation counts, from largest to smallest.	12	12	6	5	5	4	4	4	2	2
Step 2	Look up the 10% SNIP threshold for the Publication Year	1	2	3	4	5	6	7	8	9	10
Question:	How do I calculate *h*-index										
Answer:	Is the citation count for the publication (step 1) as least as large as its rank (step 2)?	12 ≥ 1 Yes	12 ≥ 2 Yes	6 ≥ 3 Yes	5 ≥ 4 Yes	5 ≥ 5 Yes	4 ≥ 6 No	4 ≥ 7 No	4 ≥ 8 No	2 ≥ 9 No	2 ≥ 10 No
	Identify the highest value for which citation count is at least as large as the publication's rank.	*h*-index ≥ 1	*h*-index ≥ 2	*h*-index ≥ 3	*h*-index ≥ 4	*h*-index = 5	*h*-index = 6	*h*-index = 7	*h*-index = 8	*h*-index = 9	*h*-index = 10
Question:	How do I calculate *g*-index?	***h*-index = 5**									
Answer:	Sum the citation counts of the publications ranked up to and including the current position.	12 12	12+12 24	12+12+6 30	12+12+6+5 35	12+12+6+ 5+5 40	12+12+6+ 5+4 44	12+12+6+ 5+5+4+4 48	12+12+6+ 5+5+4+4+4 52	12+12+6+5+ 5+4+4+4+2 54	12+12+6+5+ 5+4+4+4+2+2 56
	Calculate the square of the rank (from step 2).	1x1 1	2x2 4	3x3 9	4x4 16	5x5 25	6x6 36	7x7 49	8x8 64	9x9 81	10x10 100
	Is the sum of the citation counts at least as large as the sqaure of the rank?	12 ≥ 1 Yes	24 ≥ 4 Yes	30 ≥ 9 Yes	35 ≥ 16 Yes	40 ≥ 25 Yes	44 ≥ 36 Yes	48 ≥ 49 No	52 ≥ 64 No	54 ≥ 81 No	56 ≥ 100 No
	Identify the highest value for which the sum of the citation counts is at least as large as the square of the rank.	*g*-index ≥ 1	*g*-index ≥ 2	*g*-index ≥ 3	*g*-index ≥ 4	*g*-index ≥ 5	*g*-index ≥ 6	*g*-index ≥ 7	*g*-index ≥ 8	*g*-index ≥ 9	*g*-index ≥ 10
Question:	How do I calcuate *m*-index?	***g*-index = 6**									
Answer:	Earliest Publication Year in set	1997									
	Current Year	2014									
	Calculate number of years since earliest publication.	2014 – 1997 + 1 18									
	Divide *h*-index by number of years	5 / 18 **m-index = 0.3**									

The *h5*-index is the *h*-index based upon data from the last 5 years. For example, when selecting 2017, the *h5*-index will reflect the date range 2013-2017 for both documents and citations.

Research Intelligence

Elsevier's Research Intelligence solutions answer the most pressing challenges researchers and research managers face, with innovative solutions that improve an institution's and individual's ability to establish, execute and evaluate research strategy and performance.

For more information visit
https://www.elsevier.com/research-intelligence

Elsevier offices

ASIA AND AUSTRALIA
Tel: + 65 6349 0222

JAPAN
Tel: + 81 3 5561 5034

KOREA AND TAIWAN
Tel: +82 2 6714 3000

EUROPE, MIDDLE EAST AND AFRICA
Tel: +31 20 485 3767

NORTH AMERICA, CENTRAL AMERICA AND CANADA
Tel: +1 888 615 4500

SOUTH AMERICA
Tel: +55 21 3970 9300