

Research Intelligence

SciVal

Usage and Patent Metrics Guidebook



SciVal offers quick, easy access to the research performance of over 12,000 research institutions and 230 nations and regions—so you can visualize research performance, benchmark relative to peers, develop collaborative partnerships and analyze research trends.



ELSEVIER

Contents

1. Usage as a data source

1.1 The importance of research intelligence based on multiple data sources	6
1.2 Why are usage metrics valuable?.....	7
1.2.1 Optimal decisions recognise that research is multi-faceted.....	7
1.2.2 Usage metrics are complementary to other types of research metrics.....	7
1.2.3 Viewing activity can occur as soon as research is available online	7
1.2.4 Usage reflects engagement of the whole research community.....	7
1.2.5 Not all research is published with the expectation of being cited.....	7
1.3 What if different data sources give different messages?	9

2. Usage data and metrics

2.1 Usage data sources	10
2.2 Anonymized usage data	10
2.3 What do usage data mean?	11
2.4 Which usage events are included?	11

3. Selection of appropriate metrics

3.1 Clarity on the question being asked.....	12
3.2 Factors besides performance that affect the value of a metric.....	14
3.2.1 Do I need to address these factors?.....	14
3.2.2 Discipline.....	15
3.2.3 Database coverage	17
3.2.4 Manipulation.....	17

4. Usage metrics

4.1 Metric: Views Count.....	18
4.2 Metric: Views per Publication	20
4.3 Metric: Field-Weighted Views Impact.....	21

5. Patent Metrics

5.1 Patent basics	24
5.1.1 What is a patent?.....	24
5.1.2 What can be patented?.....	25
5.1.3 Patent Lifecycle.....	25
5.1.4 Where can I file for patent protection?.....	25
5.2 Analysis of patent data	26
5.2.1 What kind of information do patents contain?	26
5.2.2 What are patent citations?	26
5.3 SciVal's socio-economic metrics.....	26
5.3.1 How does SciVal use patent citations?	26
5.3.2 What is the coverage of patent data in SciVal?..	26
5.3.3 Why should I look at citations in patents?.....	26
5.3.4 What do SciVal patent-related metrics mean? ..	26
5.4 Further reading & links.....	27



Foreword

Usage data are generated when requests are made for online scholarly information. These data hold intelligence about the interest in research outputs, and are an important piece of the jigsaw puzzle that builds up to a 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

The availability of online usage data is a relatively recent phenomenon, and research metrics derived from usage data are not yet commonplace. Usage-based insights into impact are less familiar than insights based on publication and citation data, and funding awards data, and there are questions that have not yet been answered.

This Guidebook provides information about usage data and metrics to answer some of your questions, and to help you to start to include this intelligence in the picture that you build of the impact of research. But of course, using such information will stimulate more questions, and we do not have all the answers yet.

We are very much looking forward to working with you to learn about the new insights you can gain from this innovative information, and to answer some of the open questions.

I hope that you find this Guidebook useful, and perhaps even interesting.

Dr. Lisa Colledge
Elsevier

1. Usage as a data source

1.1 The importance of research intelligence based on multiple data sources

Research intelligence aims to understand, as completely as possible, an entity's impact on the world. This entity may be, for example, a single publication or a set of several publications, a researcher or a team or network, a research area, an institution or a country, or the research financed by a particular funder. Whatever the entity is, its total impact is multi-dimensional,¹ and is the combination of many different outputs and outcomes, such as productivity, frequency with which it has been read and cited, generation of intellectual property and spin-out companies that employ people in the region, and impact on society. This is illustrated in Figure 1.

Data sources alone will never be enough to tell the whole story. Human judgment is essential to supplement and interpret the intelligence that is present in the data. Understanding the total impact of research, on both the research community and on society, can be seen as a jigsaw puzzle, with quantitative inputs forming some of the pieces, and qualitative inputs forming others. All of the pieces are needed to see the complete picture, but a good impression can be gained from having several pieces in place, even if there are a few gaps.

We aim to make the quantitative section of that jigsaw puzzle as complete as possible, and to expand the range of data sources that we offer. Research encompasses many activities: publishing novel contributions, reading and citing, producing raw data and sharing it with others, collaborating within and between academia and business, building up a reputation and being considered an authority, and providing benefits to society outside the world of academia. These outputs and outcomes also act as a means of attracting talent and securing funding.

Usage data is generated by those who access electronic research publications. They visit a database of publications, and select and view information about their question or interest. These actions are captured by the database and form a data source called "usage data". Usage data is associated with different types of databases, such as commercial, institutional, and disciplinary, and these are all rich sources of intelligence about how research literature is being consumed.

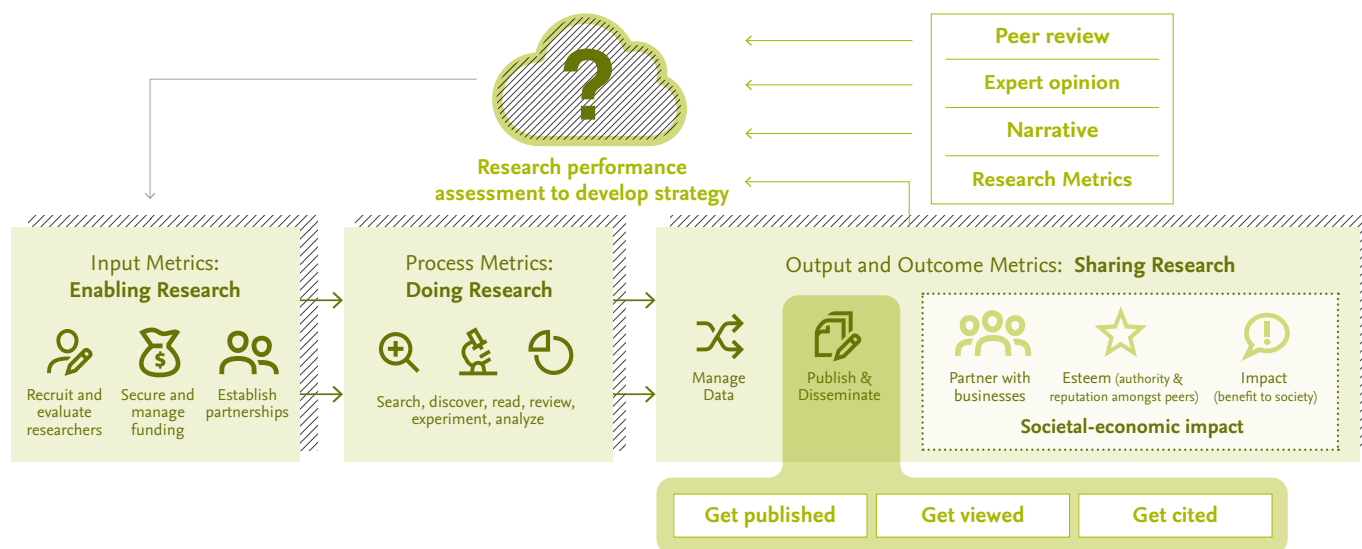


Figure 1 Research workflow. The impact of an entity, whether a single publication, a researcher, an institution, or the research financed by a particular funder, for example, is multi-dimensional, and can best be understood by combining metrics measuring a combination of inputs, processes, and outputs and outcomes. Usage data are generated by those who view electronic publications ("Get viewed"). Quantitative input should always be interpreted using judgment, and complemented by qualitative input.

1: J. Bollen, H. Van de Sompel, A. Hagberg, and R. Chute, A Principal Component Analysis of 39 Scientific Impact Measures (2009), PLoS ONE 4(6): e6022. doi:10.1371/journal.pone.0006022.

1.2 Why are usage metrics valuable?

1.2.1 Optimal decisions recognise that research is multi-faceted

There are many ways in which research can be considered excellent: for instance, it may be well cited, but it may also receive few citations and yet be well read. Optimal decisions about research performance will likely draw on the widest base of information possible, and use both qualitative and quantitative input.

Quantitative input is the most reliable when it is based on metrics that draw on a combination of data types, such as usage, publication and citation, collaboration, altmetrics, funding / awards information, and so on (see [Figure 1](#)). All of these pieces of information are complementary; usage data can reveal everything that has been viewed, and citation data represent a selection that the author of a publication has chosen to make. Combining them tells the most complete story, and can best support decision-making processes.

1.2.2 Usage metrics are complementary to other types of research metrics

The initial reaction to new research, the influence it goes on to develop, and its ultimate use by others in their own research is a complex phenomenon that cannot be adequately measured by a single criterion. Metrics drawn from multiple data sources reflect different types of behavior, with different motivations underpinning them, and all may be important in their own right.

Viewing activity, that produces usage data, is sometimes considered only in terms of how well it can predict citation activity. However, citations should not be seen as the leading research outcome against which all other behavior is to be compared. Viewing activity is important in its own right, and not only in relation to citation activity (see [Box 1](#)).

1.2.3 Viewing activity can occur as soon as research is available online

Viewing activity is typically detected before citations start to be received ([Figure 2](#)). Viewing metrics provide an early

indication of interest in an output, or set of outputs. An emerging trend, or “hot topic” in research, or a new talent, may be more quickly spotted via viewing activity, since it is more immediate than citation activity. This does not mean, however, that only research that has recently become available is viewed. An analysis of subscribed ScienceDirect full-text usage demonstrates that older publications continue to be well used.

1.2.4 Usage reflects engagement of the whole research community

Non-publishing – and hence non-citing or cited – users are estimated to constitute one-third of the research community.² This includes large numbers of undergraduate and graduate students, as well as researchers operating in the corporate sector. By incorporating demand from these users, which is “hidden” from citation-based approaches, usage-based metrics may provide a more representative indication of how research publications perform. This takes us a step closer to measuring scholarly influence on the entire research and student community.

1.2.5 Not all research is published with the expectation of being cited

Clinical research is primarily aimed at practitioners who are working with patients. These practitioners tend to read voraciously to stay up to date with new clinical advances so that they can offer their patients the best care, but they are less likely to publish original research themselves. They may eventually be cited in reviews, but this type of clinical research may be poorly cited, but very well viewed and / or read. Viewing activity may be more appropriate than citations as an indicator of impact in these disciplines.

Researchers in Arts & Humanities usually do not publish frequently, tend not to include long reference lists in publications, and their output may be of local or regional interest. The volume of citations, or citation potential, is therefore low. This should not necessarily be interpreted as the research being poor, but as a reflection of the behavior inherent in this field. Citations may not be the most useful indicator of impact in these cases, but the amount of interest in these outputs could still be significant, and this may be better measured by viewing activity based on usage data.

2: D.J.S. Price and S. Gürsey, *Studies in Scientometrics I. Transience and continuance in scientific authorship* (1976), *International Forum on Information and Documentation*, 1(2), 17-24; and C. Tenopir and W. King, *Towards electronic journals: Realities for scientists, librarians, and publishers* (2000), Washington, DC: Special Libraries Association.

Box 1: Relationship between viewing and citation activities

Viewing and citation statistics reflect different types of behavior, with different motivations underpinning them. Both are important in their own right, rather than only as predictors of each other. If there is any relationship between viewing and citation activities, it should be considered as a cycle: the influence of viewing on citation, and the influence of citation on viewing.

Table 1 shows the disciplinary correlation between full-text downloads and citations at the journal level. It is based on download counts received in the year of publication, and citation counts in the third year after publication. Correlation indicates the relationship between two sets of data, but does not necessarily mean that one causes the other; it is possible that there are one or more additional factors involved. The extent of correlation, between downloads and citations, depends

on the discipline, and, while the research did not have access to sufficient information about the user and reader populations to rigorously test the reasons for this variable correlation, the authors of this research suggest that:

- Disciplines in which the correlation is high, such as Biochemistry and Molecular Biology, tend to be specialized, and the author and the reader populations tend to coincide
- Disciplines in which the correlation is lower may have a reader population that is much broader than the publishing (cited and citing) research community. This would include readers interested in humanities and social science research from outside these disciplines, and practitioners using technical information from engineering and nursing journals

Correlation (Pearson's R)	Disciplines		
Over 0.65	Agricultural and Biological Sciences Biochemistry and Molecular Biology Business Chemical Engineering Chemistry Decision Sciences	Economics and Finance Engineering Immunology Materials Science Mathematics Medicine	Neuroscience Nursing Pharmacology Veterinary Science
0.40 to 0.65	Computer Science Dentistry Earth Sciences	Energy Environmental Science Physics and Astronomy	Psychology Social Sciences
Below 0.40	Arts and Humanities	Health Professions	

Table 1 Disciplinary correlation between full-text downloads and citation data at the journal level³.

This publication goes on to note that there is more variance in the correlation between downloads and citations at the level of individual publications. In an applied science journal, the citation counts of highly downloaded articles (>2,000 downloads) showed a strong scatter, and the journal contained highly downloaded papers which were not highly cited. Similarly, a case study of the journal *Tetrahedron Letters*⁴ found no statistical evidence of a relationship between early full-text downloads of a publication, and the citations it subsequently received. However, more of the highly cited publications than would be expected were also highly downloaded, leading to the

hypothesis that a small group of publications that were both highly downloaded and cited were responsible for driving the apparent correlation. The correlation between usage and citations is also greatly reduced when only non-English language journals are investigated.⁵

Citations also correlate with increased usage. The case study of *Tetrahedron Letters*, for instance, found that during the three months after receiving a citation, the number of full-text downloads received by a publication increased by 25% compared to what would be expected if the citation had not been received.³

3: Table 1 is based on Figure 6 in G. Halevi and H.F. Moed, Usage patterns of scientific journals and their relationship with citations (2014), *Proceedings of the science and technology indicators conference 2014 Leiden*, p241-251.

4: H.F. Moed, Statistical relationships between downloads and citations at the level of individual documents within a single journal (2005), *Journal of the American Society for Information Science and Technology*, 56(10), 1088-1097.

5: V.P. Guerrero-Bote and F. Moya-Anegón, Downloads versus citations and the role of publication language (2014), *Research Trends* 37, available at <http://www.researchtrends.com/issue-37-june-2014/downloads-versus-citations-and-the-role-of-publication-language/>



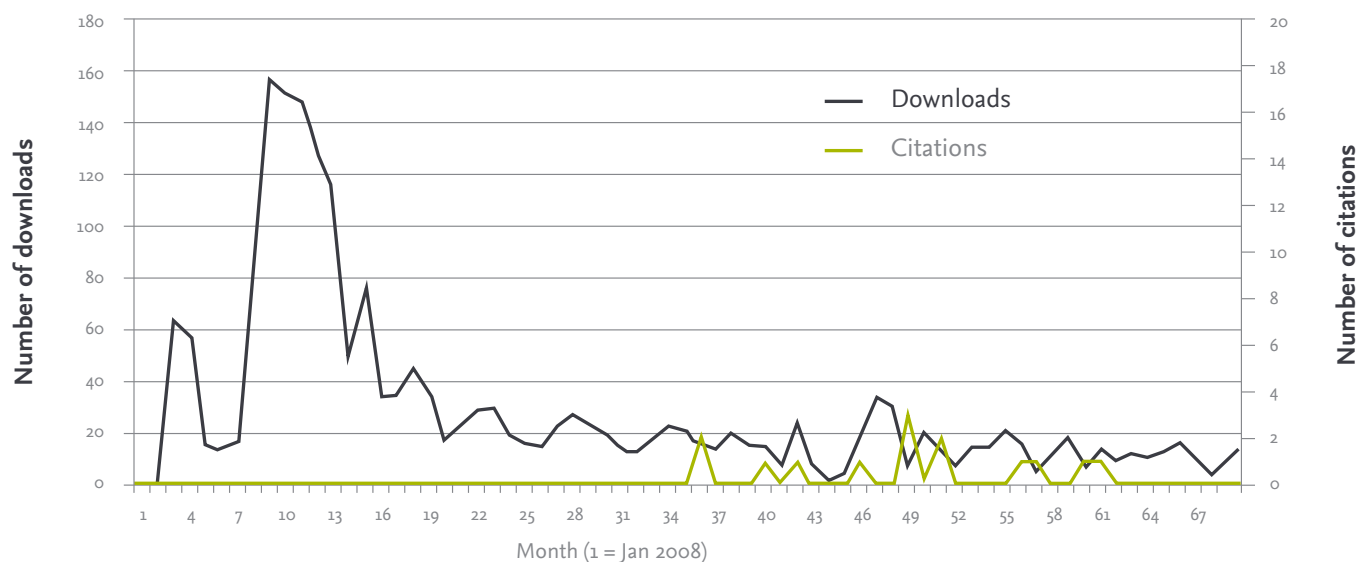


Figure 2 Usage data accumulates quickly, and is a more immediate indicator of attention than citation data.⁶ The chart shows the example of one publication whose corrected proof appeared online on 4 March 2008 (month 3 in the figure), and whose corrected paginated proof appeared online on 22 August 2008 (month 8).

1.3 What if different data sources give different messages?

The data sources underpinning research metrics are based on the day-to-day activities of researchers, students, and readers, and therefore offer useful windows into behavior and trends. Distinct metrics often reinforce each other's message, giving a high degree of confidence in the analysis and conclusions. Extra confidence due to this reinforcing effect may be especially warranted when the metrics are calculated from distinct data sources.

There will be other situations in which metrics, whether from the same or distinct data sources, appear to give conflicting information. A common reaction is that one or other data source must be incorrect, but this can be a valuable signal that further investigation would be useful. Research metrics only reflect what is present in the data produced by the research community itself, and so a more productive approach is to try to understand the reason for the apparent discrepancy. For instance:

- If there is high usage but little or no citation activity, is this because too little time has passed since publication for citations to have accumulated in this discipline? Or is this a discipline where citations are not to be expected?
- If there is citation activity but no or little usage, is this because the usage cannot be captured in the data source that is being viewed? This would be the case when using ScienceDirect usage data for publications that are not included in a journal published by Elsevier and that are therefore not available on ScienceDirect. This could be addressed by selecting Scopus usage instead.

Research metrics alone will never be enough to provide a complete picture, and human judgment and other sources of insight are essential to supplement and interpret the intelligence that is present in the data.

6: Figure 2 is based on work conducted for the following paper, and is reproduced with permission: G. Halevi and H.F. Moed, Usage patterns of scientific journals and their relationship with citations (2014), *Proceedings of the science and technology indicators conference 2014 Leiden*, p241-251.

2. Usage data and metrics



2.1 Usage data sources

The usage metrics in Elsevier's tools draw on anonymized usage data from our commercial database Scopus.⁷ All ScienceDirect and 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."⁸

Scopus is the world's largest abstract and citation database, and delivers a comprehensive overview of global research output. Scopus indexes content from over 5,000 publishers, including Elsevier, and its usage data offer the optimal representation of what is being viewed across multiple publishers. Its content is determined by the independent and international Scopus Content Selection and Advisory Board.⁹ See [Box 2](#) for more information about the content of Scopus.

2.2 Anonymized usage data

The metrics that Elsevier produces are anonymized. They do not provide any information about what a particular institution's users are viewing, and it is not possible to see the usage of a particular customer. For instance:

- SciVal¹⁰ displays information about the total views that an institution's publications have received. This can be sliced by country and sector (academic, corporate, government, or medical). This global usage includes the institution's own views, but also the views from all other institutions in the world. You can see what is being viewed, but not who is responsible for the viewing.

7: <http://www.elsevier.com/online-tools/scopus>

8: <http://www.projectcounter.org/>

9: <http://www.elsevier.com/online-tools/scopus/content-overview#scopus-content-selection-and-advisory-board>

10: <https://www.elsevier.com/solutions/scival>



ELSEVIER

Box 2: Scopus content

Scopus, the largest abstract and citation database of peer-reviewed literature, features smart tools to track, analyze and visualize research. Scopus delivers the most comprehensive overview of the world's research output in the fields of science, technology, medicine, social sciences and arts and humanities. As research becomes increasingly global, interdisciplinary and collaborative, you need to make sure that crucial research from around the world is not missed.

75+ million records

- 37+ million post-1970 records, including references (84% include abstracts)
- 6.5+ million pre-1970 records going back as far as 1788
- 8.5+ million Gold Open Access articles
- 9 million+ Conference papers

24,600+ active titles

- 23,500+ peer-reviewed journals, of which more than 4,000 are Gold Open Access
- 740+ book series¹¹
- 300+ trade publications
- Articles-in-press (i.e., articles that have been accepted for publication) from over 8,000 titles from international publishers, including Cambridge University Press, the Institute of Electrical and Electronics Engineers (IEEE)

191,000+ books

- Including monographs, edited volumes, major reference works and graduate level text books
- Focuses on social sciences and arts & humanities, but also includes science, technology & medicine (STM)

2.3 What do usage data mean?

The common feature of all types of usage activity is that a user makes a request to a service for a particular piece of scholarly information.¹² This request may be made for a variety of reasons which are unknown by the service that records the request: perhaps they are referring to a reading list, or a colleague has just published or informed them about something; perhaps they were intrigued by the title of the publication and requested it to see whether it was of interest (and it may or may not have been); or perhaps they intend to read the information and incorporate it into their own research (and may or may not eventually do so). The most that we can say is that usage reflects an interest or need for particular information.

2.4 Which usage events are included?

SciVal aims to give the most complete picture of the viewing activity within a particular research area. Therefore, it does not attempt to distinguish between different “types” of requests for information, and includes the following usage events:

- Scopus – the sum of abstract views, and clicks on the link to view full-text at the publisher's website. These events cover all views from both commercial and trial customers.

11. Scopus' journal and book title lists are available at <http://www.elsevier.com/online-tools/scopus/content-overview>

12. M. J. Kurtz, M.J., & J. Bollen, “Usage Bibliometrics” (2010) *Annual Review of Information Science and Technology* 44 (1), pp. 3-64.

3. Selection of appropriate metrics

This topic has been covered in detail in section 3 of the Research Metrics Guidebook,¹³ and is not repeated in its entirety here. Key points are detailed, together with additional information relevant to usage data.

The aim of using research metrics as an input into decision making is to complement qualitative inputs and increase confidence that the judgment is the optimal one. Elsevier offers a range of research metrics from which to select, and appropriate selection depends on two important factors:

- The question that is being asked
- Awareness 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 question being asked.

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 two 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 three 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.

¹³: <https://www.elsevier.com/research-intelligence/resource-library/research-metrics-guidebook>



3.2 Factors besides performance that affect the value of a metric

There are six factors, besides performance, that may affect the value of a metric:

- Size
- Discipline
- Publication-type
- Database coverage
- Manipulation
- Time

Discipline, database coverage and manipulation are considered further in this Guidebook. The reader is referred to the Research Metrics Guidebook for information on the remainder.

3.2.1 Do I need to address these factors?

Sometimes these factors may not need to be addressed at all, or may be used to advantage. A large institution that is aiming to present its performance favorably in order to attract students may purposefully use a metric like Views Count that does not take size into consideration. This would not, however, be a suitable approach for the evaluation of entities of varying size.

Metrics themselves may address some of these factors, as summarized in Table 2. If metrics are being used to evaluate entities of different sizes, then using the size-normalized metric Views per Publication instead of Views Count would compensate for this difference. If these entities also have a very different disciplinary profile, then Field-Weighted Views Impact might be the preferable choice.

The tools in which metrics are embedded may provide an answer, even if the metric itself does not. Views per Publication, itself sensitive to disciplinary differences, could still provide a useful in evaluating institutions with distinct disciplinary profiles if functionality is used to “slice and dice” these institutions to a common disciplinary portion.

	Size-normalized?	Field-normalized?	Publication-type-normalized?	Resistant to database coverage?	Difficult to manipulate?	Time-independent?
Views Count						
Views per Publication						
Field-Weighted Views Impact						

Table 2 Characteristics of usage metrics

3.2.2 Discipline

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

The geographical distribution of titles indexed in Scopus, is shown in [Figure 3a](#). [Figure 3b](#) shows the average Views per Publication, per discipline. Values are smaller for recent years because publications that became available online in 2018 have had less time to be viewed than those that became available in 2018, and so have accumulated fewer total counts.

3.2.3 Database coverage

For information about the coverage of Scopus, see section 2.1 and Box 2.

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 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 two 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 3a](#). 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. For more information about Scopus content coverage see Research Metrics Guidebook (3.2.4 Database coverage).

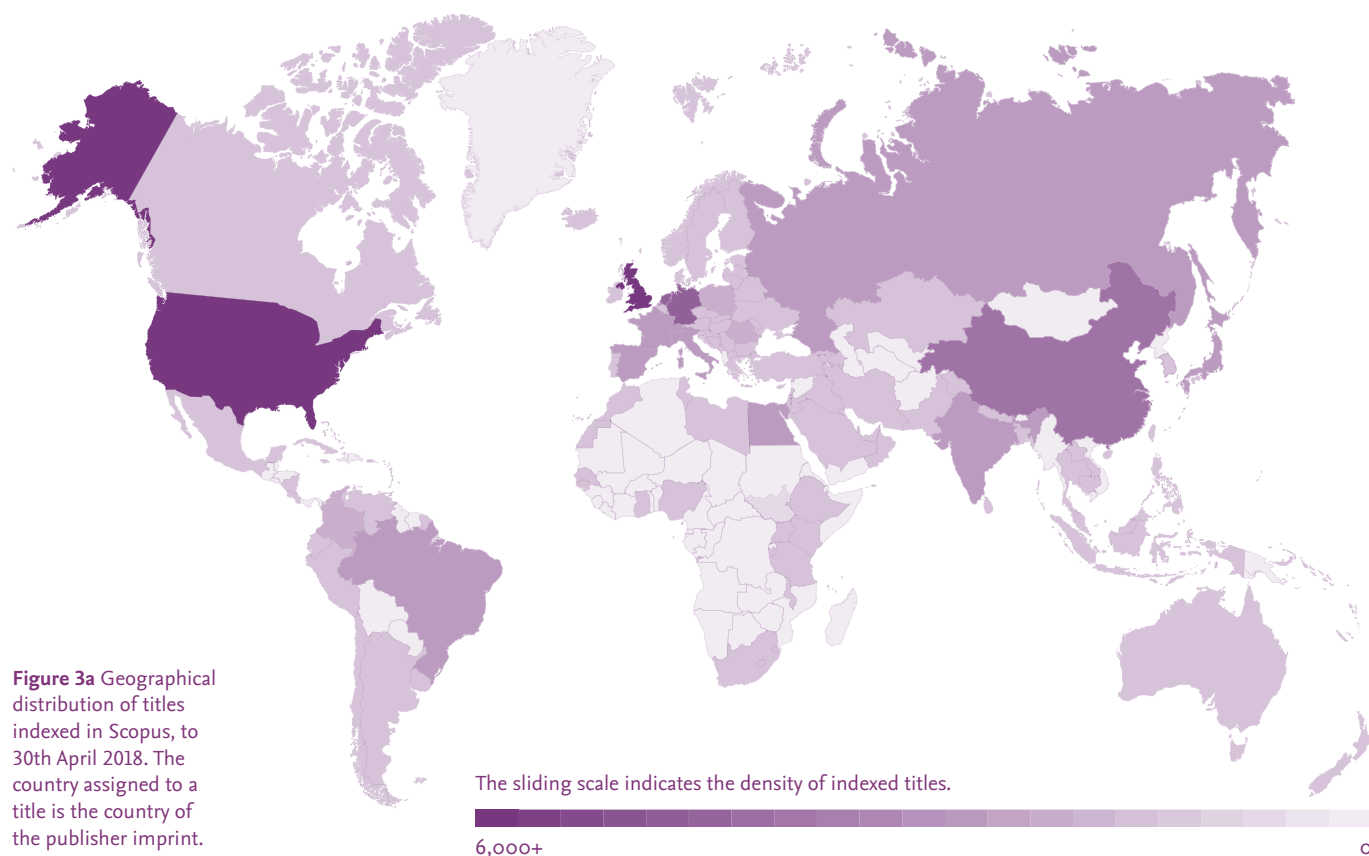


Figure 3a Geographical distribution of titles indexed in Scopus, to 30th April 2018. The country assigned to a title is the country of the publisher imprint.



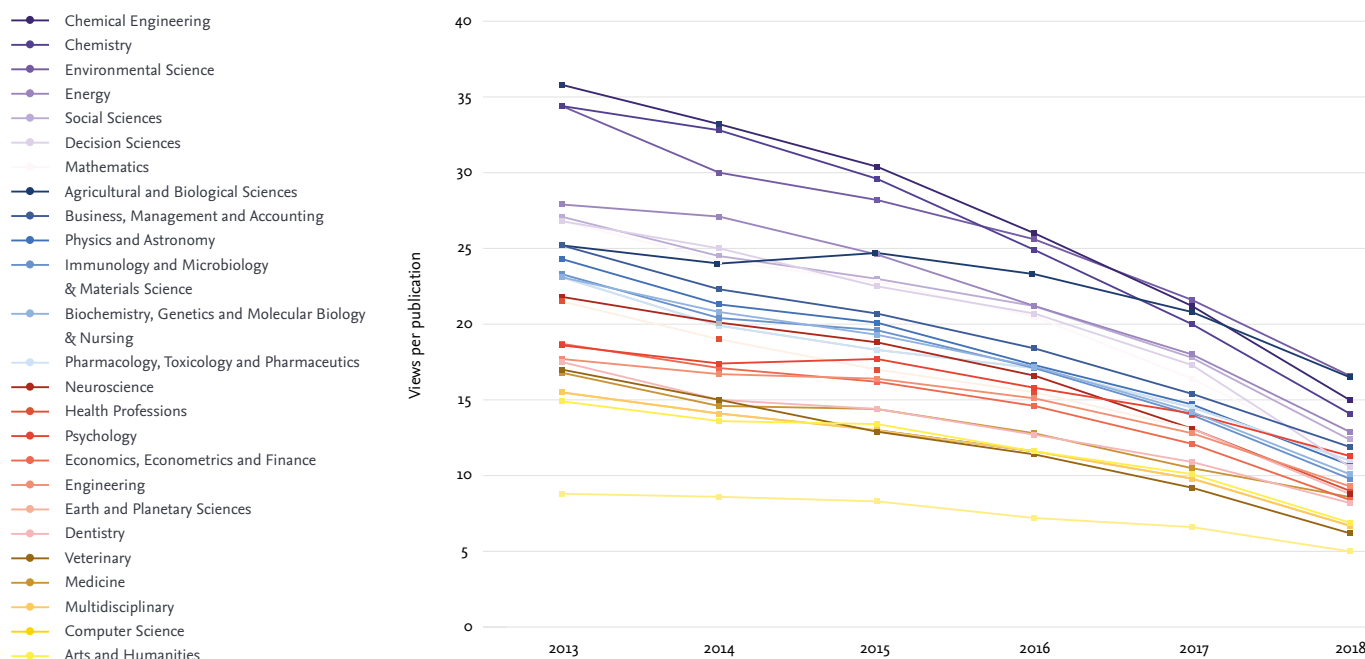


Figure 3b Shows Views per Publication per discipline calculated using Scopus usage data. Date of data cut was 15 March 2019.

3.2.4 Manipulation

Ease of manipulation of usage data is sometimes a concern in using it as an input for decision making. It is probably true that it is easier to manipulate usage data than citation data. This is another reason for drawing on multiple research metrics as input into your questions: it is much more difficult to manipulate the data underlying multiple metrics, especially if they are drawn from distinct data sets.

Moreover, there are industry guidelines in place to limit the effectiveness of attempted manipulation of usage data. Scopus usage data is COUNTER-compliant, and is 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.”¹⁴ COUNTER have prepared clear guidelines in their Code of Practice¹⁵ that address these concerns, and excerpts from this code are quoted here:

- “The intent of double-click filtering is to remove the potential of over-counting which could occur when a user clicks the same link multiple times, typically due to a slow internet connection. Double-clicks, i.e. two clicks in succession, on a link by the same user within a 30-second period MUST be counted as one action.”
- “A double-click may be triggered by a mouse-click or by pressing a refresh or back button. When two actions are made for the same URL within 30 seconds the first request MUST be removed and the second retained.”
- “For the purposes of COUNTER, the time window for a double-click on any page is set at a maximum of 30 seconds between the first and second mouse clicks. For example, a click at 10:01:00 and a second click at 10:01:29 would be considered a double-click (one action); a click at 10:01:00 and a second click at 10:01:35 would count as two separate single clicks (two actions).”

14: <http://www.projectcounter.org/>

15: <https://www.projectcounter.org/code-of-practice-five-sections/abstract/>

4. Usage metrics

This section covers the usage metrics that are currently available from Elsevier. It shares their method of calculation, situations in which they are useful, and situations in which care should be taken. It also suggests usage metrics that might be considered useful partners, either to address shortcomings of a particular metric, or to highlight information that is naturally complementary. These suggestions should not be taken as rules, but as guidelines that may sometimes be useful, and are summarized in Table 3.

Natural complementary partner metrics
Communicate information about magnitude of metrics values
Avoid display of viewing “dip” in recent years



		... useful partner metrics are:		
		Views Count	Views per Publication	Field-Weighted Views Impact
For these metrics...	Views Count			
	Views per Publication			
	Field-Weighted Views Impact			

Table 3 Suggested partner metrics

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 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 years show 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.
- In My Research Dashboard, the months and years show when a publication was 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 undergraduate 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.
- Provide transparency on the underlying data to build trust in research metrics.

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 (see [Figure 3b](#)), 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 where 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.

Views Count is calculated analogously to Citation Count. For a worked example of the calculation underlying this metric, please see Example 3 in the Research Metrics Guidebook.¹⁶

16: <https://www.elsevier.com/research-intelligence/resource-library/research-metrics-guidebook>, pages 59

4.2 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 a:

- Usage Impact metric

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

- In SciVal, the years show 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 natural complement to Views per Publication and takes into account behavioral differences between disciplines avoids the “dip” in recent years due to the most recent publications having had less time to receive views than older ones.

Views per Publication is calculated analogously to Citations per Publication. For a worked example of the calculation underlying this metric, please see Example 3 in the Research Metrics Guidebook.¹⁷

17: <https://www.elsevier.com/research-intelligence/resource-library/research-metrics-guidebook>, pages 59

4.3 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, of either 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 viewed 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% less cited 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 years show 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 takes into account 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.

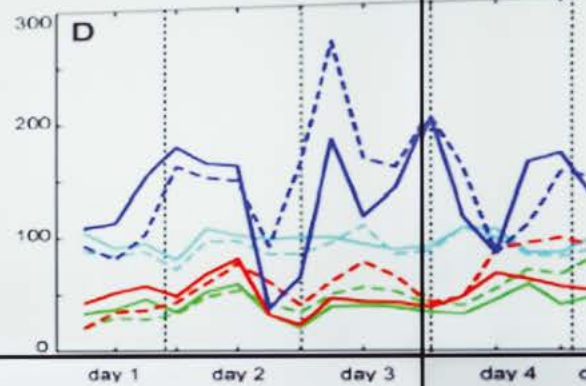
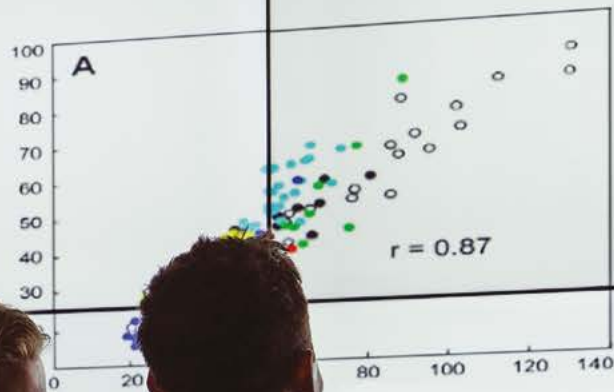
Useful partner metrics are:

- 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.

Field-Weighted Views Impact is calculated analogously to Field-Weighted Citation Impact. For a worked example of the calculation underlying the metric, please see Example 5 in the Research Metrics Guidebook.¹⁸

For the mathematical notation of this metric, please see page 63 of the Research Metrics Guidebook.

¹⁸: <https://www.elsevier.com/research-intelligence/resource-library/research-metrics-guidebook>, pages 61.



5. Patent metrics

Executive summary

- Patents protect technical inventions, which are new and can be applied in industry.
- A full technical description of the invention must be disclosed in a patent application.
- All patent information lays in the public domain and is therefore fully available.
- It takes around 18 months for a patent application to be published, which means a time-lag of approximately 18 months before the data is available to be used in SciVal.
- Patent protection can be attained in any country but is normally subject to strategic or business purposes due to costs that patent maintenance incurs.
- Patents, in the same way as scientific publications, also contain references to previous work done within the same field.
- Patents usually cite other related patents or scholarly output (scientific publications).
- SciVal looks at the citations of scholarly output in patents and provides links to both the citing patents and cited Scopus articles. This helps showcase connections between science and industry as well as the knowledge flows.
- SciVal covers patents from 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).
- SciVal's patent-related metrics serve as additional tools to demonstrate research impact.

5.1 Patent basics

Highlights

- Patents protect technical inventions from commercial exploitation by 3rd parties for a limited time-frame and in the countries where the patent protection has been attained.
- In order to be patentable, inventions must be novel, inventive and industrially applicable.
- Full technical description of an invention, disclosed in the application, will be publicly available worldwide and not limited to the countries where the patent protection will be attained.
- Patents often cite research papers along with other patents.

5.1.1 What is a patent?

A patent is an exclusive right granted for an invention, which is a new technical solution to a problem. Technical information about the invention must be disclosed to the public in a patent application.¹⁹ Patent protection gives its owner the right to prevent 3rd parties from exploiting the patented invention for commercial purposes within 20 years from the date of application.

Being publicly available, patent information is therefore an important source of technical knowledge, as well as an essential element of statistical analysis to observe innovation, technology trends and R&D activities at regional, country, institution and individual levels.

¹⁹: <http://www.wipo.int/patents/en>



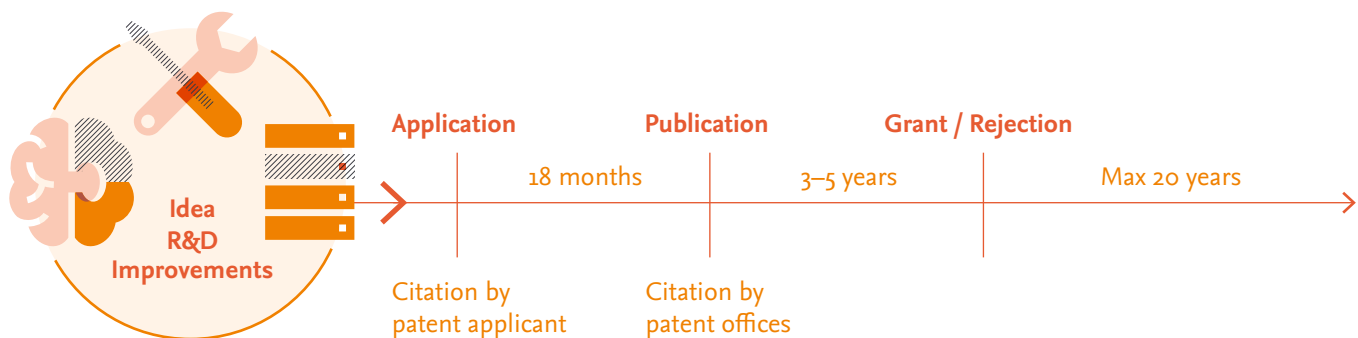


Figure 4 Patent lifecycle and majority of patent citations

5.1.2 What can be patented?

Inventions need to meet the following requirements:

- Novelty – it should not form part of existing state of the art, and should not have been published before the application.
- Inventiveness – it should not be obvious to a skilled person within the state of the art.
- Industrial application – the research should lend itself to industrial application.

More information:

UK Government Intellectual Property Office –
Protecting your patent abroad
<https://www.gov.uk/government/publications/protecting-your-uk-intellectual-property-abroad/protecting-your-patent-abroad>

5.1.3 Patent Lifecycle

All patent information is publicly available and can be found in patent databases. However, it takes around 18 months for a patent application to be published after the initial application date. Therefore there is a time-lag in the availability of patent information – everything we see today is at least 18 months old. It takes a further 3 to 5 years for a patent application to be granted or rejected by a patent office.

5.1.4 Where can I file for patent protection?

Patent protection can be attained in any and all countries worldwide. A researcher can submit patents in multiple countries, however, due to incurring filing and maintenance costs, research institutions normally restrict their country selection to core strategic countries and markets. Decisions about whether to grant a patent is taken by the individual patent offices, and guidelines may differ from country to country.

5.2 Analysis of patent data

Highlights

- Patent documents contain information about inventors, owners, countries where the invention is protected, technical field of invention and references to both patent literature and scholarly outputs.
- Patent data is publicly available and is used for statistical analysis to measure innovation, technology trends, and R&D activities, among others.
- Patent citations to scholarly output indicate a connection between research and industry, with original research as an input into innovation.
- Patents often cite research papers along with other patents.

5.2.1 What kind of information do patents contain?

Each patent contains a technical description of an invention, applicant (owner) details, inventor names and references to both patent literature and scholarly output (research papers). Additionally, each patent is classified according to patent subject classification systems.

Historically academia collaborates far more with itself than with industry, so patents and their citations provide a relevant collaboration indicator between industry and basic research.

5.2.2 What are patent citations?

Patents, as well as research literature, contain references to previous work that gives background information about an invention. These citations may be to other patents, or to original research publications. The way patent citations differ from citations in research papers is that patent citations are provided not only by the applicant (“author”) but also by the patent examiners who are reviewing the applications in the various country patent offices.

There is a difference from citations in the research world is that patent citations are provided not only by the applicant but also by the examiners who are reviewing the application in the various country patent offices.

More information: Organisation for Economic Co-operation and Development OECD

<https://www.oecd-ilibrary.org/>

5.3 SciVal’s socio-economic metrics

Highlights

- SciVal looks at citations from patents to scholarly output.
- 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.

5.3.1 How does SciVal use patent citations?

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 18 months after the application date.

5.3.2 What is the coverage of patent data in SciVal?

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).

5.3.3 Why should I look at citations in patents?

Citations from patents to scholarly outputs indicate a link between academia and industry, in other words knowledge flows. It is not possible from patents to see whether the results of the research are eventually commercially exploited, but research cited by patents is a strong indicator of the relevance that research could have to industry.

5.3.4 What do SciVal patent-related metrics mean?

Along with other indicators available in SciVal, patent-related metrics are intended to be used alongside qualitative input to showcase research impact.

- Citing-Patent Count – This is the count of patents citing the scholarly output published by the entity (e.g. a university) that you are looking at. i.e. 200 patents have cited articles published by Athena University over the past 5 years.



ELSEVIER

5.4 Further reading & links

- 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. Click “View list of patents” so see a list of the citing patents. Drill down to the patent abstract and underlying data for additional insights into this metric.
- Patent-Cited Scholarly Output – This is the count of scholarly output published by the entity (e.g. a university) that have been cited in patents. i.e. 400 publications from Athena University have been cited by patents. Click “View list of publications” to see a list of the cited scholarly output in Scopus. Drill down to the article abstract and underlying data for additional insights into this metric.
- Patent-Citations Count – This is the total count of patent citations received by the entity (e.g. a university). i.e. Athena University has been cited 600 times by patents over the past 5 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.
- 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. divide the patent-citation counts by the total scholarly output of the university for that period of time and multiply by 1,000. So if Athena University had published 10,000 publications in the 5 year period, their patent-citations per scholarly output would be $(600/10,000) \times 1,000 = 60$. We look at this metric per 1,000 publications because otherwise the typical average patent citations per output is a small number and harder to interpret.

How to apply for a European patent
<https://www.epo.org/applying/basics.html>

IPR Helpdesk – free of charge advice regarding intellectual property provided to EU funded research projects and EU SMEs
<http://www.iprhelpdesk.eu/>

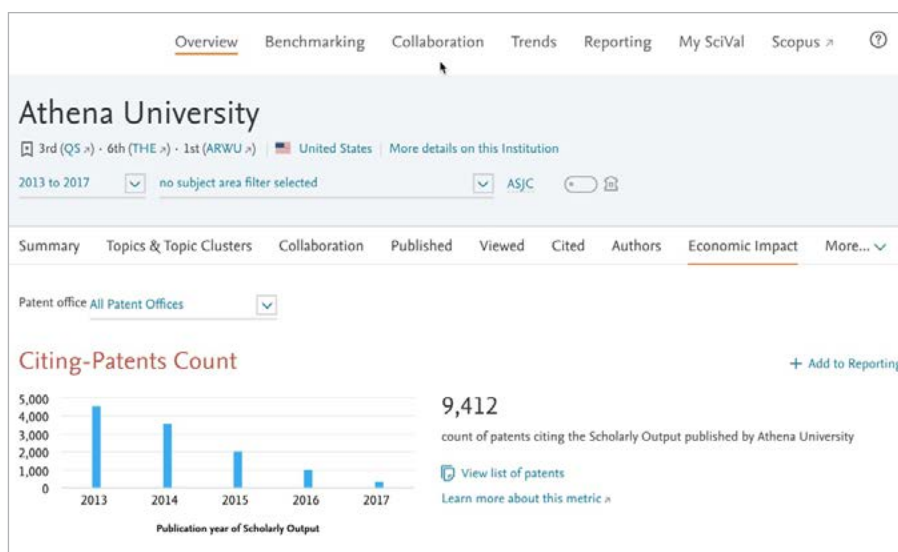


Figure 5 For all of the charts, the year on the x-axis relates to the year of the scholarly output that the patents have cited, not the year that the patents were published. Example analysis taken from the Overview module in SciVal



For more information about SciVal,
visit elsevier.com/scival

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

