



# **Snowball Metrics** Recipe Book

THEIR APPLICATION IN THE UNITED KINGDOM



NOVEMBER 2012



# **Snowball Metrics**

# Contents

About Snowball Metrics and this Recipe Book 7

## **INTRODUCTION 9**

Why Snowball Metrics? 11

Snowball Metrics – an agreed institutional perspective 13

The Snowball Metrics Landscape 15

Snowball Metrics in this recipe book 17

This recipe book, and next steps 18

## **OVERARCHING CONSIDERATIONS 19**

Definition and display of Snowball Metrics 21

Primary data sources, and implications for benchmarking 21

The definition of an institution 21

Counting 22

Output filters 23

UK application 23

*Researcher 23*

*FTE (full-time equivalent) count 23*

*HESA cost centre 24*

*Financial year 24*

*Funder type denominator 25*

## **INPUT METRICS 27**

Applications Volume ❄ 29

Awards Volume ❄ 31

## **PROCESS METRICS 37**

Income Volume ❄ 39

Market Share ❄ 41

## **OUTPUT METRICS 47**

Scholarly Output ❄ 49

Citation Count ❄ 51

*h*-index ❄ 53

Field-Weighted Citation Impact ❄ 55

Outputs in Top Percentiles ❄ 59

Collaboration ❄ 61





Snowball Metrics are designed to facilitate cross-institutional benchmarking globally by ensuring that research management information can be compared with confidence.

In agreeing a set of robust and consistent definitions for tried-and-tested Snowball Metrics across the entire spectrum of research activities, higher education institutions are establishing a trusted and reliable foundation for benchmarking and evidence-based strategic decision making.





## About Snowball Metrics and this Recipe Book

**This Recipe Book details the agreed and tested methodologies for the first set of Snowball Metrics. It's purpose is to enable any interested party to generate these metrics in a standard, commonly understood way that supports cross-institutional benchmarking, both nationally and internationally.**

The goal of Snowball Metrics is for the sector to share its knowledge and experiences in order to build best practice in evidence-based institutional strategic planning. The approach is to agree a means to measure activities across the entire spectrum of research, at multiple levels of granularity: the Snowball Metrics Framework.

### **Snowball Metrics**

- **Are global standards that enable cross-institutional benchmarking**
- **Are defined and agreed by higher education institutions themselves, not imposed by organisations with potentially distinct aims**
- **Aspire to cover the entire spectrum of research activities in support of institutional decision making**
- **Have tested methodologies that are freely available and can be generated by any organisation**
- **Are independent of data-source**

Agreed and tested Snowball Metrics will be published openly for use by any organisation, whether for public service or commercial purposes. These metrics are not intended to replace the use of existing means of informing institutional strategy, but rather to complement them by providing a perspective that may lead to valuable new insights. Metrics are widely recognised indicators of productivity, but they are nevertheless proxies for the intensity and impact of research.

The project partners are committed to driving the adoption of Snowball Metrics within the sector. We are partnering with euroCRIS<sup>1</sup>, a not-for-profit organisation that is dedicated to the development of Research Information Systems and their interoperability, to express the Snowball Metrics in CERIF. CERIF is a freely available global standard data format which enables different systems to communicate with each other in this common language: “translating” Snowball Metrics into this standard language will greatly facilitate their adoption by the CERIF-speaking community. We are delighted that euroCRIS supported the launch of this Recipe Book at their November 2012 membership meeting.



*“I am looking forward to collaborating with the Snowball Metrics team to apply our CERIF data standard. This will make the generation and adoption of these metrics by research information systems much easier, and facilitate the sharing of metrics for benchmarking between systems from different suppliers.”*

PROFESSOR KEITH JEFFERY, PRESIDENT EUROCRIS AND  
DIRECTOR OF IT AND INTERNATIONAL STRATEGY OF STFC  
(SCIENCE AND TECHNOLOGY FACILITIES COUNCIL)

Consensus on the “recipes” for this first set of Snowball Metrics has been reached by a group of UK higher education institutions. We expect that they will apply equally well to all UK institutions. We hope to receive input from the sector, both within the UK and elsewhere, as to the broader applicability of these methods: this input will guide the development of additional national “flavours” of existing recipes, as well as entirely new Snowball Metrics, to further support national and global benchmarking.



*“I am delighted by how the team has built on the key recommendations that the sector made in the JISC report on research information management. This is a huge step towards addressing the need for an agreed national framework for data and metric standards. It’s gratifying that this partnership has meant that institutions could describe their needs directly to a supplier who could test them in a way that meets the institutions’ aims.”*

DR JOHN GREEN, CHAIRMAN OF SNOWBALL METRICS STEERING  
COMMITTEE

1 [www.eurocris.org](http://www.eurocris.org)



# Introduction

**This section addresses the origin of Snowball Metrics, their purpose, the institutional research landscape across which it is aspired to have agreed measurements, as well as the use of Snowball Metrics in this Recipe Book, and next steps.**



## Why Snowball Metrics?

The efforts and experiences of English higher education institutions in implementing research management systems were surveyed in 2010. The study also aimed to identify problems with the current approaches used, and to highlight best practice.



*“This recipe book is the best practice for how data can support institutional decision making, and we hope that making it freely available to everyone who is interested will be a valuable way to further share this knowledge. We are looking forward to input from the sector to understand the extent to which this best practice is instructive, and how it can be improved.”*

PROFESSOR MICHAEL ARTHUR, VICE-CHANCELLOR, UNIVERSITY OF LEEDS

All of the key recommendations are detailed in the box on the following page. In summary, the study revealed that institutions and funders recognise data as an essential element in strategic management and decision making, but that there is a lack of consensus on those metrics that should be used for measurement and evaluation. It showed that, without clearly defined and shared metrics, institutions find it almost impossible to benchmark meaningfully, and that this hampers their ability to establish and monitor strategic direction. It also mitigates against suppliers developing a deeper understanding of research management and developing data systems that could be used across the sector, as each stakeholder focuses on different data structures and metrics, often reflected by complex bespoke implementations.



*“It’s unique for universities to join forces and take time to consider how we would like to look at our performance, and by extension how we would like our performance to be viewed. We are pleased that all the project partners have committed to openly sharing their metric definitions; we hope that the free availability of the “recipes” will encourage enthusiastic support and adoption by universities across the UK, in Europe and further afield.”*

PROFESSOR ERIC THOMAS, VICE-CHANCELLOR, UNIVERSITY OF BRISTOL

It also emerged that the focus on, and pressures of reacting to, data requests from external bodies have meant that institutions have allowed the demands of other stakeholders to determine the data and aspects of performance that they collect and

measure, rather than considering what would be best suited to their own purposes. Moreover, external data collection tends to be undertaken by stakeholders in isolation; this means that the same data are often collected several times in different formats, which compounds the challenge of making meaningful comparisons.

**Key recommendations of the report “Research Information Management: developing tools to inform the management of research and translating existing good practice”<sup>2</sup>**

- A national framework for data and metric standards should be developed with stakeholders, and used across the sector
- Institutions should work more collaboratively with each other to harmonise their approach to research management processes, and so to reduce inefficiencies
- Institutions should develop stronger relationships with suppliers and work with them to define their needs more clearly
- Suppliers should participate in the development of data and metric standards to develop deeper understanding of needs, and to drive consistency in research systems
- Institutions and funders should work together more collaboratively to identify commonality in systems and processes, so they might share data in more cost-effective and less resource-intensive ways
- Institutions, supported by funding organisations, should be encouraged to develop long-term system strategies focused upon core research

Stakeholders were keen to address the issues and opportunities for improvement identified by this study. Thus, a subset of the participants initiated Snowball Metrics to build on these recommendations: for institutions to agree definitions of metrics as a reliable base for benchmarking input into their strategic decisions.

---

<sup>2</sup> Research Information Management: developing tools to inform the management of research and translating existing good practice. [www.snowballmetrics.com/wp-content/uploads/research-information-management.pdf](http://www.snowballmetrics.com/wp-content/uploads/research-information-management.pdf)

## Snowball Metrics - an agreed institutional perspective

The goal of Snowball Metrics is for the higher education sector to share its knowledge and experiences to build best practice in evidence-based institutional strategic planning. The approach is to agree a means to measure activities across the entire spectrum of research, at multiple levels of granularity: the Snowball Metrics Framework.

*“Universities frequently struggle to locate reliable benchmarking information and Snowball Metrics provide a starting point from which to develop meaningful comparisons that inform decision-making. It is pleasing that the project builds upon existing standards within the sector and includes a commitment to openly share the metric framework for benefit across the UK and beyond.”*

JAMES MCELNAY, PRO-VICE-CHANCELLOR (RESEARCH AND POSTGRADUATES), QUEEN'S UNIVERSITY BELFAST

Snowball Metrics will ensure that robust, commonly understood and defined methodologies can be applied, at an appropriate and agreed level of detail, to research management information across the sector, whether by institutions, funding bodies or suppliers. This consistency facilitates benchmarking between institutions, helping to establish a trusted and reliable foundation for strategic decision making.

The universities, which agreed to collaborate with Elsevier, and which collectively account for approaching 40% of competitive funding awarded by Research Councils, UK-authored articles, and UK citations, are:

- University of Oxford
- University College London
- University of Cambridge
- Imperial College London
- University of Bristol
- University of Leeds
- Queen's University Belfast
- University of St Andrews

Elsevier has supported these institutions in taking forward Snowball Metrics. The institutional project partners and Elsevier have all contributed expertise and staff time to the initiative. Institutional representatives are typically directors of research offices or senior staff responsible for research management and strategy, as well as technical specialists with expertise in data systems, structure and use. All parties have provided their resources to the project for free.

## The Snowball Metrics Landscape

The institutional project partners first defined the landscape of research activities to be included in the scope of Snowball Metrics. It is intended that the landscape lends itself both to the more challenging questions of how to derive meaningful information in social sciences and arts and humanities, as well as to the somewhat more familiar approaches in STEM<sup>3</sup> disciplines. Consequently, the landscape includes the inputs, process, and outputs of institutional research activities, which are further distinguished as those connected to research grants, postgraduate education, and enterprise activities.

A set of denominators was also defined, with two functions:

- Denominators fractionalise (“slice and dice”) Snowball Metrics at various levels. The intention is to generate Snowball Metrics at those degrees of granularity appropriate to support strategic decision making, but not necessarily to use these denominators exhaustively: for example, denominators such as category of academic staff, institution, disciplinary units, and subject themes, give strategic insight across the institution
- Normalise across institutions, to enable comparison between units of different sizes

The Snowball Metrics Landscape is shown in Figure 1.



*“We are pleased that a supplier has bought in to the aspirations of this group of universities, and helped us to establish this initiative. There is a strong need for an agreed set of metrics across universities, that don’t rely on a particular supplier to calculate them, so that we can understand our strengths in relation to our peers.”*

PROFESSOR GUY ORPEN, PRO VICE-CHANCELLOR (RESEARCH),  
UNIVERSITY OF BRISTOL

---

<sup>3</sup> STEM: Science, Technology, Engineering and Mathematics

*Figure 1: The Snowball Metrics Landscape*

	<i>Research Inputs</i>	<i>Research Process</i>	<i>Research Outputs</i>
<b>Research Grants</b>	<b>Research applications</b> <b>Research awards</b> <ul style="list-style-type: none"> <li>• Price / overhead recovery</li> <li>• Philanthropy</li> </ul>	<b>Research income</b> <ul style="list-style-type: none"> <li>• Space utilisation</li> <li>• Staff recruitment</li> <li>• Start/ end date slippage</li> </ul>	<b>Publications &amp; citations</b> <b>Collaboration (co-authorship)</b> <b>Esteem measures</b> <ul style="list-style-type: none"> <li>• Socio-economic impact</li> </ul>
<b>Post-Graduate Education</b>	<b>PGR volumes</b> <ul style="list-style-type: none"> <li>• PGT volumes</li> <li>• International PGT volumes</li> <li>• UG to PG conversion rates</li> </ul>	<b>Post-graduate experience</b> <ul style="list-style-type: none"> <li>• Contact time</li> <li>• Facilities</li> </ul>	<b>Completion rates</b> <ul style="list-style-type: none"> <li>• Alumni / destination of leavers</li> <li>• Skills development (impact)</li> </ul>
<b>Enterprise Activities</b>	<b>Industrial income</b> <ul style="list-style-type: none"> <li>• Industry engagement</li> </ul>	<ul style="list-style-type: none"> <li>• Contract turnaround times</li> <li>• Industry research income</li> </ul>	<b>Patenting</b> <b>Licensing income</b> <b>Spin-out generation / income</b> <ul style="list-style-type: none"> <li>• KTPs numbers</li> <li>• Consultancy income</li> </ul>
<b>Denominators</b> <ul style="list-style-type: none"> <li>• “Slice and dice”</li> <li>• Normalise for size</li> </ul>	<b>(Number of) People</b> <ul style="list-style-type: none"> <li>• Researcher, authors</li> <li>• Principal / co-investigators</li> <li>• Academic staff by category</li> <li>• Research assistants</li> <li>• PGR Students</li> <li>• UG / PGT Students</li> <li>• Post doctoral staff</li> <li>• Support staff</li> </ul>	<b>Organisations</b> <ul style="list-style-type: none"> <li>• Institution</li> <li>• Faculty / department</li> <li>• Cost Centre</li> <li>• Unit of Assessment</li> <li>• Groups / clusters</li> <li>• Funders by type</li> <li>• Centres / institutes</li> </ul>	<b>Themes / Schemes</b> <ul style="list-style-type: none"> <li>• Standard grants</li> <li>• Strategic initiatives (Calls)</li> <li>• Grand Challenges</li> <li>• Subject areas</li> <li>• Keywords</li> </ul>

## Snowball Metrics in this recipe book

The institutional project partners have now reached a methodological consensus for an initial group of Snowball Metrics. These metrics were selected because they are:

- Familiar, and it was anticipated that reaching a consensus methodology would be relatively straightforward
- Representative of input, process and output activities
- Dependent on drawing data from each of institutional (e.g. grant applications and awards), proprietary (e.g. bibliometric data) and third party (e.g. funder) sources



*“From an information management perspective, Snowball Metrics give universities a much needed sector-designed and owned standard for benchmarking. Those involved in the design have strived to reuse existing standards in both definitions, for example categorising by HESA cost centres and HESA staff definitions, and technology, for example, working with euroCRIS to adopt and extend CERIF as the data exchange mechanism. By working with all stakeholders to promote the use of Snowball Metrics the overall aim is to provide robust benchmarking data and tools whilst keeping any increase in burden on institutions to a minimum.”*

ANNA CLEMENTS, ENTERPRISE ARCHITECT, UNIVERSITY OF ST ANDREWS

The definitions of these Snowball Metrics were initially agreed upon by technical specialists from each of the institutional project partners. The feasibility of these definitions was subsequently tested by some of the institutional project partners to ensure that they could be generated with a reasonable amount of effort that is not manually intensive. The findings of the testing have been integrated into the methodologies to optimise the scalability of Snowball Metrics, thereby ensuring that they can be readily updated to reflect the current status of an institution.

The project partners have undertaken to publish the agreed and tested standard methodologies, so that any organisation will be able to use the framework to generate Snowball Metrics<sup>4</sup>. None of the project partners will at any stage apply any charges for the methodologies of agreed and tested Snowball Metrics that have been developed. Any organisation is free to use these methodologies for their own purposes, whether these are public service or commercial.

---

<sup>4</sup> Snowball Metrics: Statement of Intent. Available at [www.snowballmetrics.com](http://www.snowballmetrics.com)

## This recipe book, and next steps

This booklet sets out the first set of Snowball Metrics Recipes. The methodologies include both generic and UK-focused approaches, together with data sources that might be considered, and opportunities for future development. The project partners expect that these methods will apply equally well to all UK institutions and thereby begin to support cross-institutional benchmarking.

The project partners are committed to driving the adoption of Snowball Metrics within the sector. We are partnering with euroCRIS<sup>5</sup>, a not-for-profit organisation that is dedicated to the development of Research Information Systems and their interoperability, to express the Snowball Metrics in CERIF<sup>6</sup>. CERIF is a freely available global standard data format which offers a very flexible structure to handle multiple sets of metrics. It enables different systems to communicate with each other in this common language. The new Indicators Task Group will “translate” Snowball Metrics into this standard language, greatly facilitate their adoption by the CERIF-speaking community.

The project partners very much hope to receive input from members of the higher education sector, both within the UK and elsewhere, as to:

- the broader relevance of the concept of metrics that are agreed and shared between institutions
- the extent to which the Snowball Metrics landscape covers all aspects of research activity that institutions would like to input into their strategies
- the degree to which sharing the agreed and tested methodologies contained in this recipe book moves institutions towards their ultimate goal of global benchmarking

Please use the Contact Us form available at: [www.snowballmetrics.com/contact-form/](http://www.snowballmetrics.com/contact-form/).

A mailing list is available to receive periodic updates of progress on Snowball Metrics.

Snowball Metrics are and will continue to be created and owned by higher education institutions. Elsevier is committed to supporting the need to develop and implement standard metrics globally for the support of institutional decision making and cross-institutional benchmarking; Elsevier intends to use Snowball Metrics in its tools, and would very much welcome their use by other suppliers. As such, it supports the further development of Snowball Metrics, whether new national flavours, new languages such as CERIF, or entirely new recipes. Future modifications to published recipes, national versions, and new agreed and tested Snowball Metrics will continue to be published free of charge by the Snowball Metrics project partners.

---

<sup>5</sup> [www.eurocris.org](http://www.eurocris.org)

<sup>6</sup> <http://www.eurocris.org/Index.php?page=CERIFreleases&t=1>



## Overarching Considerations

**This section covers agreed approaches that affect multiple Snowball Metrics, and should be consulted in conjunction with the individual recipes.**



## Definition and display of Snowball Metrics

A Snowball Metric is one which has been defined and agreed by higher education institutions as being useful in supporting strategic planning by enabling cross-institutional benchmarking. These metrics have tested methodologies to ensure that they can be generated with a reasonable amount of effort that is not manually intensive. These methodologies are freely available and can be used by any organisation.

A Snowball Metric is indicated by the use of this symbol  placed after the name of the metric.

## Primary data sources, and implications for benchmarking

The primary data sources listed are those that could be used to generate the generic Snowball Metrics.

Snowball Metrics can be calculated regardless of the specific data sources available; for example, Scholarly Output  could be generated using data from an institutional output repository or Current Research Information System (CRIS), Scopus, Web of Science, or Google Scholar. It is, however, important to have consistency in data sources when benchmarking between institutions to ensure that the comparisons are meaningful: it would be misleading for an institution to draw conclusions based on a comparison of its Scholarly Output  generated using Scopus with the Scholarly Output  of a peer institution generated using Web of Science.

For the Output Snowball Metrics, Institutional Output Repositories and CRIS<sup>7</sup> are mentioned. These include Atira's Pure<sup>8</sup> (now owned by Elsevier), Symplectic<sup>9</sup>, Avedas<sup>10</sup>, Converis<sup>10</sup>, Thomson Reuters' Research in View<sup>11</sup>, ePrints<sup>12</sup>, and dSpace<sup>13</sup>.

## The definition of an institution

For Snowball Metrics generated from institutional data, an institution is defined as the sum of data elements recorded in that institution's systems.

For Snowball Metrics generated from output data, an institution is defined as the sum of outputs associated with all name variants claimed by that institution.

- Hospitals and medical schools are considered part of the institution
- Companies are not considered part of the institution

---

7 CRIS: Current Research Information System

8 [www.atira.dk](http://www.atira.dk)

9 [www.symplectic.co.uk](http://www.symplectic.co.uk)

10 [www.avedas.com](http://www.avedas.com)

11 <http://researchanalytics.thomsonreuters.com/researchinview/>

12 [www.eprints.org](http://www.eprints.org)

13 [www.dspace.org](http://www.dspace.org)

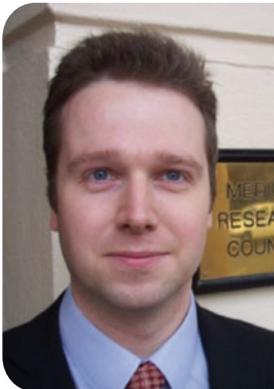
Snowball Metrics support institutional decision making, and as such data are viewed from an institutional stand-point. When a researcher moves away from an institution, the data associated with the researcher is taken as remaining with the institution: a publication generated while at institution A remains attributed to institution A even after its author has moved to institution B. A researcher's data generated while at an institution other than the one for which metrics are being considered are not included in the calculation.

## Counting

Whole counting is used to generate Snowball Metrics. The method of counting is important when a data element has more than one denominator associated with it.

For example, a data element may have multiple affiliations and researchers associated with it. Consider a publication co-authored by authors A, B and C, who are all affiliated to the same institution. Say that A and B are members of the same disciplinary denominator D<sub>1</sub>, and C is a member of a separate disciplinary denominator D<sub>2</sub>:

- In whole counting, the publication is counted as 1 publication for each denominator to give full credit to each. In this example, 1 publication will be credited to D<sub>1</sub>, and 1 publication will also be credited to D<sub>2</sub>, when reading the metric out at this level. (Fractional counting would credit both D<sub>1</sub> and D<sub>2</sub> with half a publication each)
- The data element will be deduplicated in aggregated denominators to avoid double counting. In this example, this publication will be counted once only at institutional level, despite appearing as 1 publication in each D<sub>1</sub> and D<sub>2</sub> and so counted twice at the disciplinary denominator level



*“We are very supportive of this initiative for institutions to agree standard metrics. We have experience of the significant challenges that result from data stored differently across multiple systems, and equally we are starting to experience the enormous benefits of efficiency and clarity that come from focusing on a single approach. We are excited to see how we can start to work with the Snowball Metrics project partners to streamline our approaches.”*

DR IAN VINEY, HEAD OF STRATEGIC EVALUATION, MEDICAL RESEARCH COUNCIL

## Output filters

Snowball Metrics that depend on publication output can be filtered to look at a particular subset of the total. The filtering is applied to characteristics that are associated directly with the outputs, such as the year in which they were published, or the affiliation of the authors.

Some Snowball Metrics use citation counts. Citation counts are typically lifetime citations received since publication up to the date of the current data extract. The only exception is Field-Weighted Citation Impact, which applies a 3-year citation window; for example, for an item published in October 2007, citations will be counted until the end of December 2010.

## UK application

It is a fundamental aim of Snowball Metrics that they are independent of data source. The UK application of these metrics has made use of some data from the Higher Education Statistics Agency (HESA) to facilitate consistency and ease of generation; data are returned annually to HESA by institutions, from their various systems, according to specific definitions. Further national applications may well make use of equivalent domestic data sets.

## Researcher

A researcher is defined as any employee whose contract of employment, as defined by the HESA Academic Employment Function field, is classified as either “2: Research-only” or “3: Teaching and research”, and has an Activity code of “2A: Academic Professional”.

## FTE (full-time equivalent) count

The FTE count of those Researchers returned by institutions to HESA is used wherever normalisation is needed to take into account varying sizes of institutions and disciplines. This count is updated annually.

*“Robustly defined and agreed metrics which enable funders and institutions to understand the relative strengths of the places that we fund are important for evidence-based decision making.”*

DR LIZ ALLEN, EVALUATION TEAM, WELLCOME TRUST

## HESA cost centre

HESA defines a “cost centre” as a grouping of student, staff and finance records that allows meaningful comparisons between different types of data at a more granular level than the total institution.

HESA cost centres are used in exactly this light in the context of Snowball Metrics. In identifying a discipline-level denominator that would support cross-institutional benchmarking, the institutional project partners searched for:

- A structure that has the same meaning at all UK institutions
- Information that was already available in institutional systems
- Information that was reasonably current

HESA cost centres satisfy all of these requirements. They are used as a framework for all institutions throughout the UK to return data annually to HESA. In addition, HESA cost centres do not reflect an institution’s own organisational structure or strategic priorities, unlike their departments and the Units of Assessment used by the Research Excellence Framework exercise, making them a suitable structure to support benchmarking between institutions.

*“We have been very interested to watch the progress of this initiative, and to see how institutions come together to share their experiences and knowledge, and to agree how they would like to be measured as well as providing the opportunity for benchmarking.”*

ALISON ALLDEN, CHIEF EXECUTIVE HIGHER EDUCATION STATISTICS AGENCY (HESA)

A researcher may be assigned to up to 3 HESA cost centres, although this option is applied to a very small number of researchers in the UK. The field CCENTRE1 only is used to create this denominator for Snowball Metrics. The HESA cost centre structure will change for the August 2012-July 2013 reporting year, in an effort to align data collection with that for the Research Excellence Framework’s Units of Assessment structure. HESA cost centre codes are available at: [http://www.hesa.ac.uk/dox/circulars/Cost\\_Centres/HESA\\_mapping\\_of\\_old\\_Cost\\_Centres\\_to\\_new\\_Cost\\_Centres\\_to\\_REF\\_UoAs.pdf](http://www.hesa.ac.uk/dox/circulars/Cost_Centres/HESA_mapping_of_old_Cost_Centres_to_new_Cost_Centres_to_REF_UoAs.pdf).

## Financial year

The UK higher education financial year runs from 1 August to 31 July.

## Funder type denominator

This denominator is used for Submissions Volume 🌟, Awards Volume 🌟, Income Volume 🌟 and Market Share 🌟.

Snowball Metrics denominator	Constituent HESA Funder Types	Further Breakdown for Snowball Metrics
Research Councils	<ul style="list-style-type: none"> <li>Research Councils, Royal Society &amp; British Academy</li> </ul>	<ul style="list-style-type: none"> <li>AHRC</li> <li>BBSRC</li> <li>EPSRC</li> <li>ESRC</li> <li>MRC</li> <li>NERC</li> <li>STFC</li> <li>British Academy</li> <li>Royal Society</li> </ul>
UK Charity	<ul style="list-style-type: none"> <li>UK-based Charity (QR Eligible for Charities Support)</li> <li>UK-based Charity (NOT QR eligible)</li> </ul>	
UK Public Sector	<ul style="list-style-type: none"> <li>UK central government bodies/ local authorities, health &amp; hospital authorities</li> </ul>	
UK Industry	<ul style="list-style-type: none"> <li>UK industry, commerce &amp; public corporations</li> </ul>	
Overseas Industry	<ul style="list-style-type: none"> <li>EU industry, commerce &amp; public corporations</li> <li>Non-EU industry, commerce &amp; public corporations</li> </ul>	
EU Government	<ul style="list-style-type: none"> <li>EU government bodies</li> </ul>	<ul style="list-style-type: none"> <li>European Commission</li> <li>Other EU government bodies (in aggregate)</li> </ul>
Other Overseas Sources	<ul style="list-style-type: none"> <li>EU-based charities (QR Eligible for Charities Support)</li> <li>Non-EU-based Charity (QR Eligible for Charity Support)</li> <li>EU other</li> <li>Non-EU other</li> </ul>	
Other Sources	<ul style="list-style-type: none"> <li>Other sources</li> </ul>	





## Input Metrics

**This section details the methodologies for the following Snowball Metrics:**

- **Applications Volume**
- **Awards Volume**



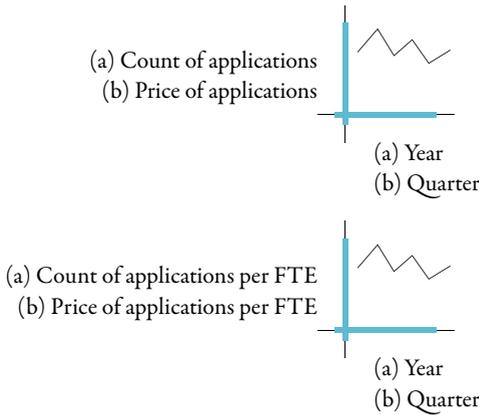


# Applications Volume<sup>\*</sup>

VOLUME OF RESEARCH GRANT APPLICATIONS SUBMITTED TO EXTERNAL FUNDING BODIES

## Metric definition

Applications Volume<sup>\*</sup> calculates the number and price of research grant applications that are submitted to external funding bodies.



## UK application

Denominators:

- HESA cost centre, via prorated mapping of departments to HESA cost centres. This mapping is done on the basis of the HESA cost centre assignment of the application's principal investigator
- Funder type (page 25)
- Institution

Calculate from 2007/2008 financial year to most recent complete quarter.

The applications considered are those that reflect activities where the resultant spend would be returned as research grants and contracts income in the HESA financial return [http://www.hesa.ac.uk/index.php?option=com\\_content&task=view&id=1145&Itemid=233](http://www.hesa.ac.uk/index.php?option=com_content&task=view&id=1145&Itemid=233).

This excludes, for example:

- Any research funding that would be passed to a collaborating institution
- Any activity that would not be considered eligible for HESA reporting, such as training activities like Doctoral Training Centres / Grants / Awards, EU Partner elements

Currency is British pounds (GBP).

## Details

The price of a research grant application is the value that the institution requests of the funder and that the funder should be willing to pay the institution to undertake the research. The price is not necessarily the same as the Full Economic Cost (fEC) to the institution to undertake the research.

Applications Volume ❄️ addresses new applications only. It excludes prior submissions in a multi-stage application process such as outlines and expressions of interest.

The date used is the date that the application is submitted to the funding body.

## Primary data source

- Institutional research grant application system

## Future opportunities

Prorated mapping of departments to HESA cost centres, on the basis of the assignment of the principal investigator, has been agreed in the UK application of this Snowball Metric. In some cases, the institutional research grant application system captures the co-investigators as well as the principal investigator. The principal investigator approach was agreed since it is inclusive and ensures that everyone can use the same methodology. It is an interesting opportunity for the future to consider mapping the department according to co-investigators as well as the principal investigator.

A denominator reflecting themes and subject focus of the competitive funding applications would be highly valued. Most likely, an automated way of assigning subject fields based on abstracts of the submissions would be needed to enable this. A critical mass of national funding bodies might be considered a source of data for this metric.

Applications Volume ❄️ may not lend itself easily to global benchmarking, due to distinct national characteristics of competitive funding structures. It might best be viewed as a metric with multiple national flavours.

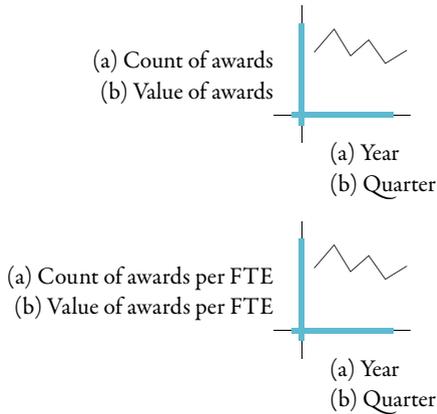


# Awards Volume

VOLUME OF AWARDS GRANTED AND AVAILABLE TO BE SPENT

## Metric definition

Awards Volume  calculates the number and value of awards from external funding bodies.



## UK application

Denominators:

- HESA cost centre, via prorated mapping of departments to HESA cost centres. This mapping is done on the basis of the HESA cost centre assignment of the award's principal investigator
- Funder type (page 25)
- Institution

Calculate from 2007/2008 financial year to most recent complete quarter.

Awards considered reflect activities where the resultant spend would be returned as research grants and contracts income in the HESA financial return [http://www.hesa.ac.uk/index.php?option=com\\_content&task=view&id=1145&Itemid=233](http://www.hesa.ac.uk/index.php?option=com_content&task=view&id=1145&Itemid=233).

This excludes, for example:

- Any research funding that would be passed to a collaborating institution
- Any activity that would not be considered eligible for HESA reporting, such as training activities like Doctoral Training Centres / Grants / Awards, EU Partner elements

Currency is British pounds (GBP).

## Details

Awards Volume ✨ considers aggregated values of awards over the award lifetime: the total value awarded at the time of award and not the value (to be) spent in any financial year.

This metric includes subsequent financial amendments to awards, including supplements and reductions, and funding from industry. It does not include non-financial amendments such as no-cost extensions.

Amendments to the value of the original award, whether positive or negative, should be treated as new awards. For example:

- A £1m award is received in 2007/2008
- If this award is increased by £0.5m in 2011, an award of £0.5m, not £1.5m, is recorded in 2010/2011
- If the award is then reduced by £0.2m in 2012, an award of -£0.2m, not £1.3m, is recorded in 2011/2012

Income received from a spin out company acting as a funder of research to the university is included in Award Volume ✨. However, any funding that a spin-out company receives, as a separate entity to the university, is not included.

The date used is the date that the award is entered in the institutional grants system. This date was selected for pragmatic reasons since it is always available, and ensures that all awards are included. The preferred date of award notification is not consistently available, and would result in less comparable values.

## Primary data source

- Institutional grants system

## Future opportunities

Prorated mapping of departments to HESA cost centres, on the basis of the assignment of the principal investigator, has been agreed in the UK application of this Snowball Metric. In some cases, the institutional research grant application system captures the co-investigators as well as the principal investigator. The principal investigator approach was agreed since it is inclusive and ensures that everyone can use the same methodology. It is an interesting opportunity for the future to consider mapping the department according to co-investigators as well as the principal investigator.

A denominator reflecting themes and subject focus of the competitive funding applications would be highly valued. Most likely, an automated way of assigning subject fields based on summaries of the awarded grants would be needed to enable this.

A critical mass of national funding bodies might be considered a source of data for this metric.

Awards Volume ❄️ may not lend itself easily to global benchmarking, due to distinct national characteristics of competitive funding structures. It might best be viewed as a metric with multiple national flavours.





## NOTES

A series of horizontal dashed lines for writing notes, spanning the width of the page.





## Process Metrics

**This section details the methodologies for the following Snowball Metrics:**

- **Income Volume**
- **Market Share**



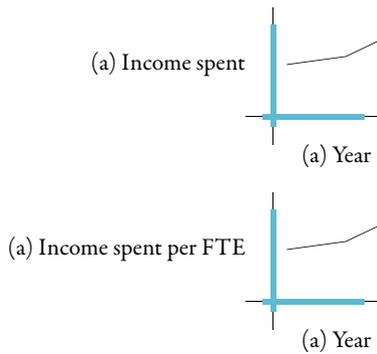


# Income Volume

VOLUME OF RESEARCH INCOME SPENT

## Metric definition

Income Volume  calculates the value of awarded budget derived from research awards from external funding bodies that has been spent.



## UK application

Denominators:

- HESA cost centre
- Funder type (page 25)
- Institution

Calculate from 2007/2008 financial year to most recent complete financial year.

Income data available from HESA Finance Record are used to generate Income Volume .

Currency is British pounds (GBP).

## Details

### Primary data sources

- Institutional accounts system
- Published annual accounts
- National statutory reports, such as those available from HESA in the UK

### Future opportunities

A denominator reflecting themes and subject focus of income would be highly valued. As yet, no suitable structure has been identified. Most likely, an automated way of assigning subject fields based on summaries of the awarded grants linked to the income would be needed to enable this.

Income Volume  may not lend itself easily to global benchmarking, due to distinct national characteristics of competitive funding structures. It might best be viewed as a metric with multiple national flavours.

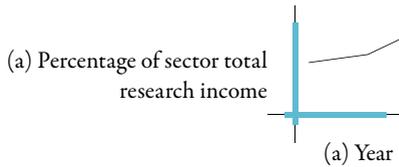


## Market Share

PERCENTAGE OF SECTOR TOTAL RESEARCH INCOME PER INSTITUTION

### Metric definition

Market Share  calculates the percentage of total research income across the sector related to a given institution.



### UK application

Denominators:

- HESA cost centre
- Funder type (page 25)
- Institution

Calculate from 2007/2008 financial year to most recent complete financial year.

Income data available from HESA Finance Record are used to generate Market Share . The sector total research income is the total national income as reported to HESA.

Currency is British pounds (GBP).

## Details

### Primary data sources

- Institutional accounts system
- Published annual accounts
- National statutory reports, such as those available from HESA in the UK

### Future opportunities

A denominator reflecting themes and subject focus of income would be highly valued. As yet, no suitable structure has been identified. Most likely, an automated way of assigning subject fields based on summaries of the awarded grants linked to the income would be needed to enable this.

Market Share  may not lend itself easily to global benchmarking, due to distinct national characteristics of competitive funding structures. It might best be viewed as a metric with multiple national flavours. Versions might be derived based on:

- Amounts awarded by funding bodies, rather than spend
- The total available amongst participating institutions, rather than the national total



## NOTES

A series of horizontal dashed lines for writing notes, spanning the width of the page.





## NOTES

A series of horizontal dashed lines for writing notes, spanning the width of the page.





## Output Metrics

This section details the methodologies for the following Snowball Metrics:

- Scholarly Output
- Citation Count
- *h*-index
- Field-Weighted Citation Impact
- Outputs in Top Percentiles
- Collaboration



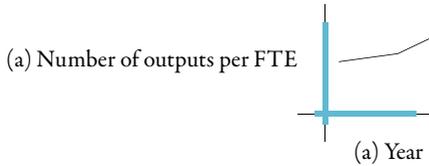
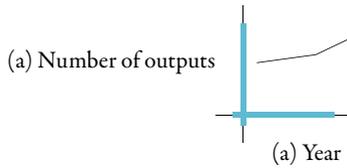


# Scholarly Output

PRODUCTIVITY BASED ON ANY TYPE OF SCHOLARLY OUTPUT

## Metric definition

The metric counts the number of institutional outputs of any type.



## UK application

Denominators:

- HESA cost centre, via assignment of a researcher associated with an output to a HESA cost centre
- Institution

Calculate from 2008 to most recent complete calendar year.

## Details

### Primary data sources

- Institutional output repository and CRIS
- Scopus
- Web of Science
- Google Scholar

When the data source is an institutional output repository or CRIS, various distinct types of output may be included in the count for Scholarly Output , such as data sets, exhibitions, and working papers. Institutions should be specific about the output types they are including in their count when benchmarking, to ensure comparability.

### Future opportunities

Commercial abstracting and indexing databases extend their degree of coverage of an institution's output to give a more comprehensive picture of an institution's activity.

A standard list of output types may be generated, including not only the typical groupings of outputs from STEM<sup>14</sup> subjects such as articles, reviews, conference papers, and editorials, but additional agreed categories such as compositions, exhibitions, working papers, and PhD theses.

The HESA cost centre denominator can be replaced by any assignment of researchers to groupings that are shared across institutions nationally and internationally. Examples would be the grouping of researchers into clusters for national evaluation exercises.

A denominator that is a subject classification covering the entire data source will support global benchmarking.

---

<sup>14</sup> STEM: Science, Technology, Engineering and Mathematics

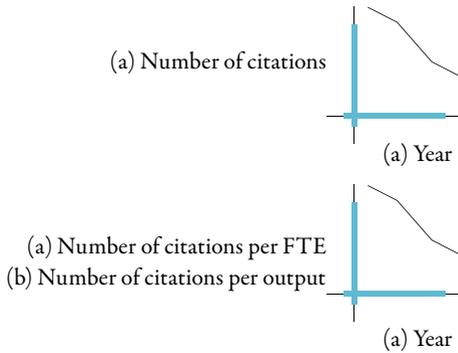


# Citation Count

CITATIONS RECEIVED BY INSTITUTION'S SCHOLARLY OUTPUT

## Metric definition

The metric counts the citations received to date by institutional outputs.



## UK application

Denominators:

- HESA cost centre, via assignment of a researcher associated with an output to a HESA cost centre
- Institution

Calculate from 2008 to most recent complete calendar year.

## Details

It is likely that citation data will not be available for all elements that constitute an institution's Scholarly Output ❄️. For example, if a commercial abstracting and indexing database is used as the data source for Citation Count ❄️, their coverage will be less than 100% of the institution's total productivity. A partial reflection of an institution's activity is still valuable in providing an evidence-based support for decision making, but the limitation should be borne in mind.

## Primary data sources

- Scopus
- Web of Science
- Google Scholar

## Future opportunities

Commercial abstracting and indexing databases extend their degree of coverage of an institution's output to give a more comprehensive picture of an institution's activity.

The HESA cost centre denominator can be replaced by any assignment of researchers to groupings that are shared across institutions nationally and internationally. Examples would be the grouping of researchers into clusters for national evaluation exercises.

A denominator that is a subject classification covering the entire data source will support global benchmarking.



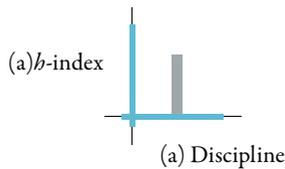
## *h*-index ❄️

H-INDEX PER DISCIPLINE

### Metric definition

The metric calculated the *h*-index, as defined by Professor Jorge Hirsch<sup>15</sup>, for institutional disciplines. To quote from this paper that defines the *h*-index in terms of researchers: “A scientist has index *h* if *h* of his or her  $N_p$  papers have at least *h* citations each and the other  $(N_p - h)$  papers have  $\leq h$  citations each”.

In other words, a group of papers has an *h*-index of 17, if 17 of these papers have each received at least 17 citations, and 18 of these papers have not each received at least 18 citations.



*h*-index ❄️ is influenced by both the quantity (Scholarly Output ❄️) and publication impact (Citation Count ❄️) of the outputs per institutional discipline.

- It can never be higher than the output regardless of that output’s impact. The *h*-index of 1 paper that has received 1,000 citations is 1
- It can never be higher than the number of citations received by the most cited paper, regardless of the amount of output. The *h*-index of 1,000 papers that have each received 1 citation is 1

### UK application

Denominator:

- HESA cost centre, via assignment of a researcher associated with an output to a HESA cost centre

Use outputs from 2008 to the current year.

15 Hirsch, J. E. (2005). “An index to quantify an individual’s scientific research output” Proc. Natl Acad. Sci. 201 (46): 16569–16572. <http://arxiv.org/abs/physics/0508025>. <http://www.pnas.org/content/102/46/16569.full>

## Details

It is likely that citation data will not be available for all elements that constitute an institution's Scholarly Output  $\otimes$ . For example, if a commercial abstracting and indexing database is used as the data source for *h*-index  $\otimes$ , their coverage will be less than 100% of the institution's total productivity. A partial reflection of an institution's activity is still valuable in providing an evidence-based support for decision making, but the limitation should be borne in mind.

## Primary data sources

- Scopus
- Web of Science
- Google Scholar

## Future opportunities

Commercial abstracting and indexing databases extend their degree of coverage of an institution's output to give a more comprehensive picture of an institution's activity.

The HESA cost centre denominator can be replaced by any assignment of researchers to groupings that are shared across institutions nationally and internationally. Examples would be the grouping of researchers into clusters for national evaluation exercises.

A denominator that is a subject classification covering the entire data source will support global benchmarking.



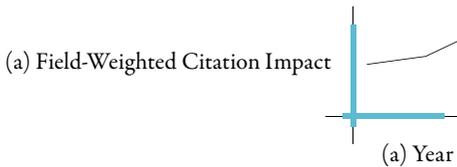
## Field-Weighted Citation Impact

ACTUAL CITATION COUNT RELATIVE TO THE EXPECTED WORLD CITATION COUNT

### Metric definition

Field-Weighted Citation Impact is the ratio of the total citations actually received by the denominator's output, and the total citations that would be expected based on the average of the subject field. A Field-Weighted Citation Impact of:

- Exactly 1 means that the output performs just as expected for the global average
- More than 1 means that the output is more cited than expected according to the global average; for example, 1.48 means 48% more cited than expected
- Less than 1 means that the output is cited less than expected according to the global average



Field-Weighted Citation Impact takes into account the differences in research behaviour across disciplines. It is particularly useful for a denominator that combines a number of different fields, although it can be applied to any denominator.

- Researchers working in fields such as medicine and biochemistry typically produce more output, with more co-authors and longer reference lists, than researchers working in fields such as mathematics and education; this is a reflection of research culture, and not performance
- In a denominator comprising multiple disciplines, the effects of outputs in medicine and biochemistry dominate the effects of those in mathematics and education
- This means that, using non-weighted metrics, an institution that is focused on medicine will appear to perform better than an institution that specialises in social sciences
- The methodology of Field-Weighted Citation Impact  accounts for these disciplinary differences

## UK application

Denominators:

- HESA cost centre, via assignment of a researcher associated with an output to a HESA cost centre
- Institution

Calculate from 2008 to most recent complete calendar year.

## Details

The expected total citation count for an output is determined based on:

- Year of publication
- Subject field
- Output type

The citations received up to 3 complete calendar years after publication are considered; for example, for an item published in October 2007, citations will be counted until end December 2010. This is an exception to the general approach of applying filters to the output item itself.

If an output is part of more than one subject field, the expected citations in each field are determined, and the harmonic average<sup>16</sup> is used as the input into Field-Weighted Citation Impact .

If an output is not assigned to a subject field, for whatever reason, then it will not be represented in the calculation.

The actual : expected ratio per output is first calculated, and then the average of these ratios is determined.

It is likely that citation data will not be available for all elements that constitute an institution's Scholarly Output . For example, if a commercial abstracting and indexing database is used as the data source for Field-Weighted Citation Impact , their coverage will be less than 100% of the institution's total productivity. A partial reflection of an institution's activity is still valuable in providing an evidence-based support for decision making, but the limitation should be borne in mind.

---

<sup>16</sup> The harmonic average is appropriate for situations when the average of ratios is desired. Definitions and examples can be found online, for example via Wikipedia.

## Primary data sources

- Scopus
- Web of Science
- Google Scholar

## Future opportunities

Commercial abstracting and indexing databases extend their degree of coverage of an institution's output to give a more comprehensive picture of an institution's activity.

The HESA cost centre denominator can be replaced by any assignment of researchers to groupings that are shared across institutions nationally and internationally. Examples would be the grouping of researchers into clusters for national evaluation exercises.

A denominator that is a subject classification covering the entire data source will support global benchmarking.



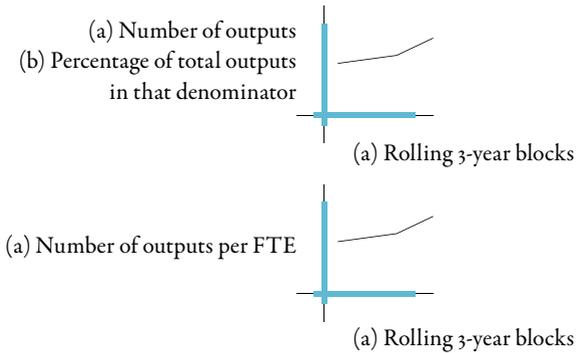


## Outputs in Top Percentiles 🚩

OUTPUTS THAT HAVE REACHED A PARTICULAR CITATION THRESHOLD IN THE DATA UNIVERSE

### Metric definition

The citation thresholds that represent the top 1%, 5%, 10% and 25% outputs in the data universe being used are established. The absolute counts, or percentage of total counts, of outputs that lie within each threshold is calculated.



### UK application

Denominator:

- HESA cost centre, via assignment of a researcher associated with an output to a HESA cost centre

Use outputs from 2008 to the current year.

## Details

It is likely that citation data will not be available for all elements that constitute an institution's Scholarly Output ❄️. For example, if a commercial abstracting and indexing database is used as the data source for Outputs in Top Percentiles ❄️, their coverage will be less than 100% of the institution's total productivity. A partial reflection of an institution's activity is still valuable in providing an evidence-based support for decision making, but the limitation should be borne in mind.

## Primary data sources

- Institutional output repository and CRIS
- Scopus
- Web of Science
- Google Scholar

## Future opportunities

Commercial abstracting and indexing databases extend their degree of coverage of an institution's output to give a more comprehensive picture of an institution's activity.

The HESA cost centre denominator can be replaced by any assignment of researchers to groupings that are shared across institutions nationally and internationally. Examples would be the grouping of researchers into clusters for national evaluation exercises.

A denominator that is a subject classification covering the entire data source will support global benchmarking.



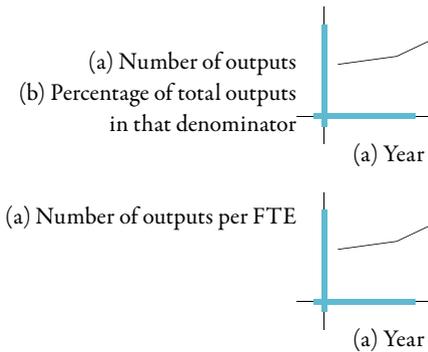
## Collaboration

### VOLUME AND PROPORTION OF NATIONALLY AND INTERNATIONALLY CO-AUTHORED SCHOLARLY OUTPUTS

#### Metric definition

Collaboration  calculates the percentage of outputs that have national or international co-authorship.

- An output has national co-authorship if it has an affiliation that does not belong to the parent institution but is within the parent institution's country
- An output has international co-authorship if it has an affiliation that does not belong to the parent institution and is outside the parent institution's country
- An output is classified as either national or international. An output that has both national and international co-authorships will be classified as international
- Countries are defined as in the ISO classification<sup>17</sup>



#### UK application

Denominators:

- HESA cost centre, via assignment of a researcher associated with an output to a HESA cost centre
- Institution

Calculate from 2008 to most recent complete calendar year.

The parent institution's country is the United Kingdom: England, Scotland, Wales and Northern Ireland.

<sup>17</sup> <http://www.iso.org/iso/ics6-en.pdf>

## Details

Institutions may have research groups or facilities affiliated to them and permanently based overseas, such as researchers in local universities, hospitals, or governmental research centres. Collaboration ❄️ considers the physical location of the affiliation's researchers to be irrelevant. As such, and taking the University of Oxford in the United Kingdom as an illustrative model:

- A collaboration between Oxford-affiliated researchers based overseas who are collaborating with an overseas institution is international
- A collaboration between Oxford-affiliated researchers based overseas who are collaborating with a UK institution is national
- A collaboration between Oxford-affiliated researchers based overseas and another UK institution's, other than Oxford, research group also based overseas is international
- A collaboration between 2 or more Oxford-affiliated researchers is an institutional collaboration, and is not included in the metric definition

The country information actually provided in the outputs is used. If an author did not include their country in their affiliation information, then their affiliation is not taken into account in the metric.

It is likely that affiliation data will not be available for all elements that constitute an institution's Scholarly Output ❄️. For example, if a commercial abstracting and indexing database is used as the data source for the collaboration information, their coverage will be less than 100% of the institution's total productivity. An institutional system may only partially capture this information for the outputs it holds. A partial reflection of an institution's activity is still valuable in providing an evidence-based support for decision making, but the limitation should be borne in mind.

## Primary data sources

Any data source that structurally captures the affiliation information of outputs, for example:

- Institutional output repository and CRIS
- Scopus
- Web of Science
- Google Scholar

## Future opportunities

Commercial abstracting and indexing databases extend their degree of coverage of an institution's output to give a more comprehensive picture of an institution's activity.

The HESA cost centre denominator can be replaced by any assignment of researchers to groupings that are shared across institutions nationally and internationally. Examples would be the grouping of researchers into clusters for national evaluation exercises.

A denominator that is a subject classification covering the entire data source will support global benchmarking.





# NOTES

A series of horizontal dashed lines for writing notes.



“From an information management perspective, Snowball Metrics give universities a much needed sector-designed and owned standard for benchmarking. Those involved in the design have strived to reuse existing standards in both definitions, for example categorising by HESA cost centres and HESA staff definitions, and technology, for example, working with euroCRIS to adopt and extend CERIF as the data exchange mechanism. By working with all stakeholders to promote the use of Snowball Metrics the overall aim is to provide robust benchmarking data and tools whilst keeping any increase in burden on institutions to a minimum.”

ANNA CLEMENTS, ENTERPRISE ARCHITECT, UNIVERSITY OF ST ANDREWS

“This recipe book is the best practice for how data can support institutional decision making, and we hope that making it freely available to everyone who is interested will be a valuable way to further share this knowledge. We are looking forward to input from the sector to understand the extent to which this best practice is instructive, and how it can be improved.”

PROFESSOR MICHAEL ARTHUR, VICE-CHANCELLOR, UNIVERSITY OF LEEDS

“It’s unique for universities to join forces and take time to consider how we would like to look at our performance, and by extension how we would like our performance to be viewed. We are pleased that all the project partners have committed to openly sharing their metric definitions; we hope that the free availability of the “recipes” will encourage enthusiastic support and adoption by universities across the UK, in Europe and further afield.”

PROFESSOR ERIC THOMAS, VICE-CHANCELLOR, UNIVERSITY OF BRISTOL

“Universities frequently struggle to locate reliable benchmarking information and Snowball Metrics provide a starting point from which to develop meaningful comparisons that inform decision-making. It is pleasing that the project builds upon existing standards within the sector and includes a commitment to openly share the metric framework for benefit across the UK and beyond.”

JAMES MCELNAY, PRO-VICE-CHANCELLOR (RESEARCH AND POSTGRADUATES), QUEEN’S UNIVERSITY BELFAST

“We are pleased that a supplier has bought in to the aspirations of this group of universities, and helped us to establish this initiative. There is a strong need for an agreed set of metrics across universities, that don’t rely on a particular supplier to calculate them, so that we can understand our strengths in relation to our peers.”

PROFESSOR GUY ORPEN, PRO VICE-CHANCELLOR (RESEARCH), UNIVERSITY OF BRISTOL



*Snowball Metrics Project Partners*

UNIVERSITY OF OXFORD  
UNIVERSITY COLLEGE LONDON  
UNIVERSITY OF CAMBRIDGE  
IMPERIAL COLLEGE LONDON  
UNIVERSITY OF BRISTOL  
UNIVERSITY OF LEEDS  
QUEEN'S UNIVERSITY BELFAST  
UNIVERSITY OF ST ANDREWS  
ELSEVIER

\* Ordered according to productivity 2011 (data source: Scopus)

[www.snowballmetrics.com](http://www.snowballmetrics.com)