# Snowball Metrics Recipe Book

Standardised research metrics – by the sector, for the sector

**Snowball Metrics Steering Group**

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

(Ordered within country according to Scholarly Output ❄, 2016.
Data source: Scopus. Data cut: 3 March 2017)

www.snowballmetrics.com

**Snowball Metrics**

Authored by Dr Lisa Colledge, Director of Research Metrics,
Elsevier, on behalf of all the Snowball Metrics program partners.

*Edition 3*

November 2017

# Contents

## Foreword: Snowball Metrics and this Recipe Book



**The aim of Snowball Metrics** is to become the international standard that enables research-intensive universities to understand their strengths and weaknesses, so that they can build and monitor effective strategies.

**Snowball Metrics is a response to common frustrations**[1,2,3] voiced by universities:

- **Informed decisions depend on data, as well as on expert opinion and peer review.** If we lack an evidence-base, we prevent ourselves from being able to make the best decisions for our universities.
- **Our systems and the data that we collect are often determined in response to frequent demands from funders and agencies.** We spend so much time collecting data in the different formats requested that we have very little opportunity to think about which systems and data would be most useful to address our own questions, which is surely what should be driving our approach.
- **Universities are poor at collaborating** with each other, and especially at working constructively with funders and agencies...

1 Research Information Management: developing tools to inform the management of research and translating existing good practice. http://www.snowballmetrics.com/wp-content/uploads/research-information-management1.pdf

2 The Current Health and Future Well-Being of the American Research University. http://www.snowballmetrics.com/wp-content/uploads/RIM_Report_Research-Futures-Consortium-.pdf

3 Moving towards the ideal way to manage research information in the United Kingdom. Report from cross-sector workshops, December 2012. https://www.snowballmetrics.com/wp-content/uploads/CrossSector_workshop_report_FINAL.pdf

- The **commercial systems and tools available have not effectively addressed all the needs of a university,** which has led to the proliferation of bespoke institutional systems. These bespoke systems have often been created independently, so that little best practice has been established; consequently commercial suppliers have struggled to determine the universal needs that they should be addressing. This negative feedback loop is not good for the higher education sector.

A general **consensus emerged in these early studies that "someone" should take ownership of finding a solution to these problems**. It was acknowledged that "someone" should lead, not only to attract attention and support to the initiative, but also to address the criticism and cynicism that would undoubtedly be encountered. Snowball Metrics was consequently initiated by a small but influential core of institutions, which together account for nearly 40% each of competitive funding awarded by the United Kingdom's Research Councils, UK-authored articles, and UK citations. The ambition is that the conclusions and approaches endorsed by this core will "infect" the international higher education sector through a "snowball effect", hence the name Snowball Metrics. These research institutions, ordered by their Scholarly Output ❄ in 2016, are:

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

The **Snowball Metrics approach** to address these problems is:
- **To enable informed evidence-based decision-making by agreeing a single method to calculate metrics that will provide input to our institutional strategies by ensuring that we are comparing apples with apples.** These metrics are based on all the data sources we have available to us, including our own institutional data sources, as well as third party and commercially available sources. **Snowball Metrics do not depend on a particular data source or supplier.**

- **To own the definition of these metrics ourselves, rather than be vulnerable to them being imposed on us by funders and agencies.** Snowball Metrics provide the opportunity to approach evaluation from our own perspective, and to think about how to best answer our own questions and needs, rather than reacting to the needs of others.
- **To collaborate with each other to agree a common solution, and to try to influence funders and agencies** to adopt this as a common solution in place of the many unique approaches that produce so much inefficiency both for institutions and funders.
- **To work together with a commercial supplier of research information** so that they can learn about our needs at first hand, and build systems and tools that enable us to effectively and efficiently store our information and provide unambiguous, rigorously defined metrics based on consistent data.

**The output of Snowball Metrics is a set of mutually agreed and tested methodologies: "recipes".** These recipes are available free-of-charge and can be used by anyone for their own purposes and, if applicable, under their own business models. I am proud to present the third edition of the Snowball Metrics Recipe Book, which offers a broad perspective on metrics that underpin effective institutional decision-making. This edition clearly demonstrates significant global progress towards our aims since publication of the first edition in November 2012.

Consensus amongst and between these stakeholders is critical to success in meeting the challenges which this initiative is addressing. **Please give your support to Snowball Metrics by championing their use within your universities, funders, agencies and suppliers.** We can transform the way in which evidence-based decisions are made across the sector, but it relies entirely on us working together and speaking with a single, unified voice across all dimensions of our influence.

*Dr John Green*
on behalf of all Snowball Metrics program partners
Chair of the Snowball Metrics Steering Group
University of Cambridge, United Kingdom

# 1. Introduction

This section covers:

- The Snowball Metrics approach: a bottom-up, academia-industry collaboration.

- Snowball Metrics recipes and their intended use.

- The pragmatic approach to getting as far as possible as quickly as possible, and efforts to reuse existing standards.

- Globalizing Snowball Metrics as standards through the free Snowball Metrics Exchange service.

- The recommended use of Snowball Metrics as a balanced scorecard for internal institutional decision-making, rather than ranking.

- Snowball Metrics case studies.

- Summary: three key points about Snowball Metrics.

## 1.1 The Snowball Metrics approach

**Snowball Metrics addresses two major needs and frustrations of research-intensive universities, initially identified in a report on the perception of institutional research management in England[4] and reiterated in a series of cross-sector workshops that were held in the UK 2 years later[5]:**
- Standard metrics to benchmark themselves against each other.
- The universities needed to be in the driving seat to truly endorse these metrics as global standards.

**Snowball Metrics is, at its heart, a bottom-up initiative.** This means that it is owned by research-intensive universities around the globe, which ensures that its outcomes are of practical use to them; its outputs are not imposed on universities by funders, agencies, or suppliers of research information.

**The aims of Snowball Metrics are:**
- **To agree on methodologies that are robustly and clearly defined so that the metrics they describe enable the confident comparison of apples with apples.** The resulting benchmarks between research-intensive universities can then be trusted as reliable information to help establish and monitor institutional strategies.
- **That the Snowball Metrics kite mark, the snowflake, becomes internationally recognized as a standard** that helps to illuminate the strengths and weaknesses of universities.
- **That institutions, funders, agencies, and suppliers of research information adopt Snowball Metrics.** A single method of requesting and consuming information will drive enormous efficiencies in all sectors of higher education, and the resources saved can be more efficiently deployed.

**Snowball Metrics is an academia-industry collaboration.** The universities involved invited Elsevier to collaborate in this initiative because they sought skills that would complement their own expertise[6]. The roles and responsibilities of the academic and industrial project partners are:
- **Everyone covers their own costs.** This is extremely important: there is no suspicion that there is a commercial agenda underpinning the conclusions of the initiative.

---

4 Research Information Management: developing tools to inform the management of research and translating existing good practice. http://www.snowballmetrics.com/wp-content/uploads/research-information-management1.pdf
This report was produced in 2010 through a collaboration between Elsevier and Imperial College London, and was sponsored by the Joint Information Systems Committee (JISC) UK.

5 Moving towards the ideal way to manage research information in the United Kingdom: cross-sector workshops. Lisa Colledge (2013). http://www.snowballmetrics.com/wp-content/uploads/CrossSector_workshop_report_FINAL.pdf

6 Initially the project partners approached all the main systems suppliers and Elsevier was the only one who chose to participate.

- **Universities:**
  - Agree which metrics will be endorsed as Snowball Metrics.
  - Determine and agree practical methodologies to generate these metrics in a commonly understood manner to enable apples-to-apples benchmarking, in a way that is independent of the different systems and data sources they have in house.
- **Elsevier:**
  - Conducts the day-to-day project management of the global program.
  - Ensures that the methodologies agreed by the universities are technically feasible, before they are shared with the global higher education sector.
  - Uses its global networks to share the outcomes, and to communicate about the initiative.

The following are **outside the remit of the Snowball Metrics program:**
- The **quality of the data sources** used to generate Snowball Metrics. These are the responsibility of the institutions, third parties, and suppliers who own these data sources.
- The **provision of tools to enable the generation and use of Snowball Metrics.** Elsevier, and any other commercial supplier, may choose to implement the recipes in their commercial offerings, but this is not part of their participation in the initiative and it is a business decision that they take independently of Snowball Metrics.

## 1.2 Snowball Metrics recipes and their intended use

The output of Snowball Metrics is a set of agreed and tested standard "recipes", or methodologies for calculating research metrics in a consistent way that enables apples-to-apples benchmarking. The recipes have been tested to ensure that they are supplier-agnostic and can be calculated from any data source that universities have available to them: in-house data sources such as applications and enterprise data, commercial data sources such as the journals and books databases available from Elsevier (Scopus[7]) and Clarivate Analytics (Web of Science[8]), and third parties such as national bodies who aggregate data from universities.

The Snowball Metrics recipes are free, and can be used by any organisation for their own purposes. The following is an extract from the publicly available Statement of Intent[9]:

---

7 https://www.elsevier.com/solutions/scopus

8 https://clarivate.com/products/web-of-science/

9 The complete Statement of Intent can be viewed at
   https://www.snowballmetrics.com/wp-content/uploads/Snowball-Metrics-Letter-of-Intent.pdf

> *"None of the project partners, including Elsevier, will at any stage apply any charges for the methodology 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."*

## 1.3 A pragmatic approach

The Snowball Metrics program takes a pragmatic approach to achieving its aims. It gets as close as possible to its objectives, and avoids the inertia that would be caused by compulsively seeking to reach the perfect endpoint, as is sometimes perceived to have been the case with the development of bibliometrics. Snowball Metrics has identified a fit-for-purpose metrics framework, recognising that this is a starting point and it will develop further over time; it has begun to source data, taking that which is readily available and trying to ensure it conforms to standards, while cognisant that something is better than nothing. This philosophy is reflected in both the route that is being taken to agree metrics to enable apples-to-apples benchmarking, and in the efficient reuse of existing standards wherever possible.

### 1.3.1 The route to agree metrics to enable apples-to-apples benchmarking
The representatives of the universities who are working on Snowball Metrics are the leaders of the research and planning offices, together with technical experts from within the universities who have experience in responding to requests for information from funders, and who consequently know the strengths and weaknesses of institutional data, systems and tools. These are people who have a daily need to use research information in order to advise, inform, and illuminate their colleagues and the external stakeholders that they work with; it is they who recognize the immediate benefit from any improvement in the arsenal of tools that they have available to them.

One of the aims of these university representatives is to **agree Snowball Metrics throughout the entire landscape of research activities** in which a research institution invests resources and would like to excel (Figure 1). In addition to agreeing the metrics themselves, a set of denominators is needed. These **denominators:**
- **"Slice and dice"** the Snowball Metrics at levels that are more granular than an entire institution, for instance to enable understanding of strengths within a discipline.
- **Normalize for size** between institutions, so that it is not always the case that bigger institutions appear to perform better.

**Figure 1: The Snowball Metrics Landscape**

| | Research Inputs | Research Processes | Research Outputs and Outcomes |
|---|---|---|---|
| **Research** | **Research applications Research awards** • Price / overhead recovery • Philanthropy | **Research income** • Space utilisation • Staff recruitment • Start / end date slippage | **Publications & citations Collaboration (co-authorship) Esteem measures** • Socio-economic impact |
| **Enterprise Activities/ Economic Development** | **Industrial income** • Industry engagement | • Industry research income | **Patenting Licensing income Spin-out generation / income** • KTPs numbers • Consultancy income |
| **Post-Graduate Education** | **PGR volumes** • PGT volumes • International PGT volumes • UG to PG conversion rates | **Post-Graduate experience** • Facilities | **Completion rates** • Alumni / destination of leavers • Skills development (impact) |
| **Denominators** • "Slice and dice" • Normalize for size | **(Number of ) People** • Researcher, authors • Principal / co-investigators • Academic staff by category • Research assistants • PGR Students • UG / PGT Students • Post-doctoral staff • Support staff | **Organizations** • Institution • Faculty / department • Cost Centre • Groups / clusters • Funders by type • Centres / institutes | **Themes / Schemes** • Standard grants • Strategic initiatives (Calls) • Grand Challenges • Subject areas • Keywords |

The project partners tackled the "low-hanging fruit" first to make progress, and then moved on to more challenging and perhaps more controversial metrics. Consequently, the first edition of this recipe book, published in November 2012, shared the agreed methods of 10 metrics that are unarguably important for institutional strategy; the focus of the project partners then was to work out how to reach consensus, and how to perform the feasibility testing of those metrics.

The second phase built on the success and progress of the first phase, and tackled some more difficult metrics. The third phase has focused on the most challenging metrics to date – Success Rate ❄, and metrics in post-graduate education and collaboration – and they are shared for the first time in this third edition of the recipe book (Figure 2).

**Figure 2: Snowball Metrics recipes**

| | Research Inputs | Research Processes | Research Outputs and Outcomes |
|---|---|---|---|
| **Research** | • Applications Volume • Awards Volume • Success Rate | • Income Volume • Market Share | **Publications & citations** • Scholarly Output (enhanced) • Citation Count • Citations per Output • h-index • Field-Weighted Citation Impact • Outputs in Top Percentiles • Publications in Top Journal Percentiles<br><br>**Collaboration** • Collaboration • Collaboration Publication Share • Collaboration Impact • Collaboration Field-Weighted Citation Impact • Academic-Corporate Collaboration • Academic-Corporate Collaboration Impact<br><br>**Societal impact** • Altmetrics • Public Engagement • Academic Recognition |
| **Enterprise Activities/ Economic Development** | • Academic-Industry Leverage • Business Consultancy Activities | • Contract Research Volume | • Intellectual Property Volume • Intellectual Property Income • Sustainable Spin-Offs (enhanced) • Spin-Off-Related Finances (enhanced) |
| **Post-Graduate Education** | • Research Student Funding | • Research Student to Academic Staff Ratio | • Time to Award of Doctoral Degree • Destination of Research Student Leavers |

Legend:
- ■ Snowball Metrics shared in original Recipe Book, November 2012
- ■ Snowball Metrics shared in edition 2 of the Recipe Book, June 2014
- ■ Snowball Metrics shared in this edition of the Recipe Book, November 2017

### 1.3.2 Reusing existing standards

**Snowball Metrics reuse existing standards whenever they support the needs of the initiative;** it does not redefine data elements and calculations that are already usefully well-defined and accepted elsewhere, but rather embraces existing work and builds upon it. Some existing standards are used exactly as they are, and some provide the basis for an adjusted definition.

One example is the support of Snowball Metrics by **the United Kingdom's Higher Education Statistics Agency** (HESA[10]). HESA's mission is to collect a comprehensive body of reliable statistics and information from the funded providers of higher education, in the areas of research, enterprise, teaching and learning, and then to provide that data back to universities and UK Research Councils. The UK Snowball Metrics Steering Group has reused and built on some of HESA's open and widely-used definitions.

A second example is the **partnership of Snowball Metrics with euroCRIS**[11], a not-for-profit organization that is dedicated to the development and inter-operability of Current Research Information Systems (CRIS systems). The euroCRIS Indicators Task Group aims to develop and share best practice in the use of indicators to support research information management. One of the outputs of the group is to express multiple sets of indicators in the euroCRIS' global data standard, the Common European Research Information Format[12] (CERIF), with Snowball Metrics being the first set to be defined. The CERIFied recipes have been used in the implementation of the Snowball Metrics Exchange service (section 1.4.1). This valuable work in partnership between Elsevier and euroCRIS highlighted some necessary enhancements to CERIF xml to improve its usability for the purposes of calculating metrics. The complete CERIF xml code for Snowball Metrics, as prepared by euroCRIS, will be available for download and use from the Snowball Metrics website[13]. The accuracy of the CERIFication of Snowball Metrics is the responsibility of euroCRIS.

Another example is the **partnership of Snowball Metrics with CASRAI**, the Consortia Advancing Standards in Research Administration and Information[14]. Through a project with CASRAI made possible by a donation from Elsevier, the Snowball Metrics project partners were able to receive input into the metric recipes from funding bodies such as The Wellcome Trust[15] and Medical Research Council[16], other suppliers including Digital Science[17] and Clarivate Analytics[18] as well as organizations such as RAND Europe[19]. The accuracy of expression of Snowball Metrics in the CASRAI data dictionary is the responsibility of CASRAI, who will make their versions of the recipes available online.

## 1.4 Globalizing Snowball Metrics as standards

The challenges being addressed by Snowball Metrics are globally relevant[20]. The Snowball Metrics initiative started in the United Kingdom because this is where the group of universities that decided to take the initial steps was located. It was always the intention to ensure that **the recipes, and the data elements that underlie them, are universal and globally relevant,** and contributions of those outside the UK continue to be apparent in this edition of the recipe book.

Importantly, this recipe book shares 2 denominators that are based on UK and US national data sources and that have been mapped to each other by UK and US Working Groups, providing practical proof of concept that international benchmarking based on locally-held data sets is possible. These 2 denominators are:

- **Discipline:** one of the UK applications of this denominator uses the HESA cost centre, and the US application is based on the National Science Foundation's Higher Education Research and Development survey (HERD) classification (section 2.9.3).
- **Funder-type:** the UK application of this denominator uses the HESA funder types, and the US application is based on the FundRef classification (section 2.13).

The project partners have connected with metrics projects in other geographies (e.g. Australia / New Zealand, United States, Asia-Pacific, Japan, Russia) which are developing national models which include metrics similar to many of the Snowball Metrics. We continue to collaborate in order to achieve a common core of metrics which can be used for global benchmarking.

### 1.4.1 The Snowball Metrics Exchange service

Snowball Metrics can be used within a single institution to give useful information about trends over time, but their real value and motivation is for benchmarking and that requires institutions to be able to see each other's Snowball Metrics. This is already possible to some extent within benchmarking tools that have implemented these recipes based on commercial data, such as Elsevier's SciVal[21], but maximum benefit depends on users being able to understand their position relative to their peers on a wider set of metrics, including those that rely on institutional data such as Applications Volume ❄.

---

10 http://www.hesa.ac.uk/
11 www.eurocris.org
12 http://www.eurocris.org/Index.php?page=featuresCERIF&t=1
13 https://www.snowballmetrics.com/
14 http://casrai.org/
15 https://wellcome.ac.uk/
16 https://www.mrc.ac.uk/
17 https://www.digital-science.com/
18 https://clarivate.com/
19 https://www.rand.org/randeurope.html

20 The current health and future well-being of the American research university. A report from the Research Universities Futures Consortium. https://www.snowballmetrics.com/wp-content/uploads/RIM_Report_Research-Futures-Consortium-.pdf
21 https://www.elsevier.com/solutions/scival

This need has given rise to the **Snowball Metrics Exchange service**[22], a free "broker service", which has been built by Elsevier such that:
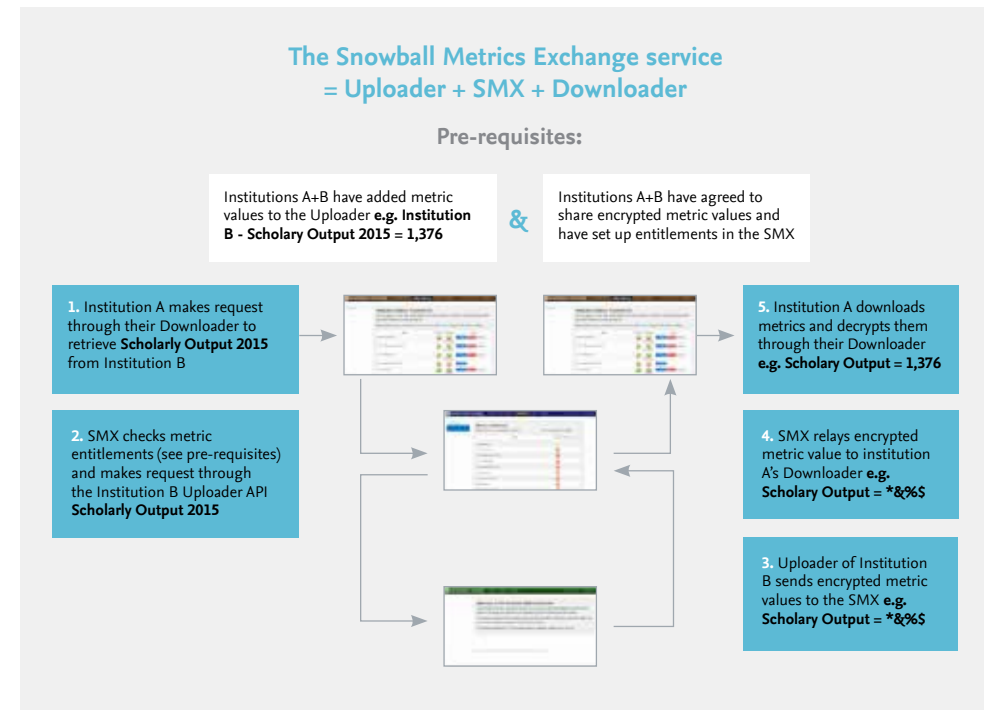
- Any institution who is using Snowball Metrics can become a member of the Snowball Metrics Exchange service.
- The institutional members will be responsible for generating their Snowball Metrics according to the recipes. The metrics could be calculated using a bespoke system, in a spreadsheet, or in a commercial tool.
- Each institution can be a member of one or more **benchmarking clubs**: groups of institutions which have agreed to exchange metrics with each other.
- Institutions may choose to accept or decline requests to share all or some Snowball Metrics with benchmarking clubs or individual institutions; this is entirely under their control.
- Institutions will use the **"I'll show you mine if you show me yours"** facility in order to exchange equivalent Snowball Metrics with each other.
- Only Snowball Metrics values will be exchanged. The data underlying the metrics will never be exchanged, and will remain behind the institutions firewalls.

The Snowball Metrics Exchange service is made of three components (Figure 3):

- The Snowball Metrics Uploader ("Uploader") that includes an API and allows an institution to encrypt and upload Snowball Metrics from their own institution for future exchange.
- The Snowball Metrics Exchange ("SMX"), which acts as the 'broker' to relay the encrypted metrics from an institution's Uploader for download by entitled institutions. Within the SMX, an institution can manage the metric entitlements for each institution that has agreed to share metrics with them.
- The Snowball Metrics Downloader ("Downloader") that allows an institution to download encrypted metrics shared by entitled, participating institutions.

Any gaps, caveats and other factors regarding the data used to generate the Snowball Metrics, that are important for the receiving university to be aware of to help in their interpretation of the benchmarking, should be shared in the notes section of the Snowball Metrics Exchange service.

**Figure 3: Overview of the flow of information within the Snowball Metrics Exchange service**



The Snowball Metrics Exchange service
= Uploader + SMX + Downloader

Pre-requisites:

Institutions A+B have added metric values to the Uploader **e.g. Institution B - Scholary Output 2015 = 1,376** **&** Institutions A+B have agreed to share encrypted metric values and have set up entitlements in the SMX

**1.** Institution A makes request through their Downloader to retrieve **Scholarly Output 2015** from Institution B

**2.** SMX checks metric entitlements (see pre-requisites) and makes request through the Institution B Uploader API **Scholarly Output 2015**

**5.** Institution A downloads metrics and decrypts them through their Downloader **e.g. Scholary Output = 1,376**

**4.** SMX relays encrypted metric value to institution A's Downloader **e.g. Scholary Output = *&%$**

**3.** Uploader of Institution B sends encrypted metric values to the SMX **e.g. Scholary Output = *&%$**

## 1.5 The recommended use of Snowball Metrics[23]

Metrics are a strong complement to qualitative inputs, such as peer review, expert opinion and case studies, when making research management decisions, and the ideal situation is to have information from quantitative as well as qualitative sources. If intelligence from these complementary approaches "triangulates", or in other words gives a similar message, then this increases the confidence in conclusions. Conflicting messages might suggest that further investigation is needed. It is also strongly advised to "triangulate" within both the quantitative and qualitative types of input, as any single source of input will not be able to address the multi-faceted nature of the question at hand.

**Snowball Metrics offers a balanced scorecard based on a broad set of metrics:**

- It is strongly advisable to "triangulate" within the quantitative input into a decision. Every metric has weaknesses, but these can be compensated for by the strengths of (an)other metric(s). It is the responsibility of the user of Snowball Metrics to ensure that a metric's weaknesses are compensated for by another metric or input.

- There is a broad diversity of questions that metrics could be used to help address. Existing scorecards often tend to be based upon output and citation metrics, largely since comprehensive commercial databases are readily available, and / or financial metrics, since they are relatively easy for universities to measure. Snowball Metrics draw a much more comprehensive and rounded view of institutional performance across the full range of research activities.

- They are unlikely to distort the research process in unanticipated ways through encouraging too much focus on a particular activity. For example, it is well known that rewarding researchers solely for publishing a high volume of output encourages researchers to slice their work in more, smaller pieces to be able to publish a higher volume[24], which is probably not a true reflection of the desired outcome. Snowball Metrics offers the option to select several metrics to encourage a balanced outcome for universities.

- Some metrics may be more or less relevant to different disciplines. Many metrics, such as Applications Volume ❄ and Awards Volume ❄, are equally useful across all fields, when the disciplinary denominator is used and an appropriate selection of peers for benchmarking is made. Other metrics, such as Citation Count ❄ and Collaboration ❄ may be more valuable in STEM[25] areas and less so in the social sciences and arts and humanities.

- Some practitioners of metrics prefer simple, straightforward metrics, often based on total counts, such as Citation Count ❄ and Altmetrics ❄. Others prefer more complex metrics, that often inherently correct for variables such as those between disciplines, for instance Field-Weighted Citation Impact ❄.

## 1.6 Snowball Metrics case studies

The **purpose of Snowball Metrics is to understand institutional strengths and weaknesses**, so that this intelligence can be used internally to inform university strategies. They help to inform universities about where they are with respect to the peer institutions they monitor, and provide information on where performance is strong and where further investment might be valuable.

Snowball Metrics are not motivated by international ranking. It is hoped that, because the driver behind Snowball Metrics is internal and not for showcasing or ranking, **there is no motivation to "game" the metrics**, because there is no gain for an institution in concealing its standing amongst its peers from itself.

Snowball Metrics should be seen as **a balanced scorecard of metrics from which a selection can be made** to help understand institutional strengths and weaknesses in a particular area. They are not intended to be prescriptive, in that one does not have to use them all at any one time nor for any one purpose – the opportunity is there to use whichever of them is felt might add value to decision making in any particular situation. The selection of Snowball Metrics will depend on the question being asked; the selection may differ from day to day even for one person, and the sub-set of the metrics used, for instance 1-5, or 2, 9, and 13, is up to the universities and funders who may well have their own preferences and opinions about the usefulness of any combination of metrics in particular circumstances. In a way, it is just like using a recipe book to cook your dinner: it is not necessary to cook the entire book to find it useful; instead, select what you like, and perhaps what you have the ingredients in the fridge (data) for.

### 1.6.1 Development of key performance indicators at the University of St Andrews[26]

Universities need robust metrics to help develop and monitor evidence-based strategies, but there is the challenge of ensuring that these metrics can be used with confidence. This is the challenge that is solved by Snowball Metrics. Consequently, when the University of St Andrews recently reviewed its institutional-level key performance indicators, it used Snowball Metrics where relevant. This allowed the university to benefit from the tried and tested 'recipes', arrived at through discussion, debate, and testing by a group of experts from leading universities, as well as positioning the university to benchmark with peer institutions and departments who are also using Snowball Metrics recipes.

Success Rate ❄ is amongst the key performance indicators being used by the University of St Andrews. The value of using Snowball Metrics recipes is well illustrated by considering the range of questions that needed to be answered before determining the method to be used to calculate this metric:

- What should be counted? Applications that are known to have been successful and unsuccessful, certainly, but how should those applications that have not yet received a definite decision ("write-off") be handled?
- After what period should an application be considered to be unsuccessful, if no decision has been received? Should this period vary for different funders?

---

24 Moed HF, Research Assessment Exercises: Informed judgments on research quality or quantity? Scientometrics 2008; 74(1): 153-161.

25 STEM: Science, Technology, Engineering and Mathematics.

26 Clements, Darroch and Green. 2017. Snowball Metrics – providing a robust methodology to inform research strategy – but do they help? Procedia Computer Science 106: 11-18, http://www.sciencedirect.com/science/article/pii/S1877050917302715

- Which amount should be considered? The original requested price or the final (negotiated) price awarded?
- In which year should an award, rejection, or write-off be counted? The application year, or the year of the award/rejection/write-off?

The Snowball Metric recipe for Success Rate ❄ answers all these questions, and represents the consolidated expertise of specialists from some of the leading universities in the UK. For example:

- Regarding which amount should be considered, the project partners confirmed that since the final price was always available in the internal systems from which the data for this metric would be drawn, whereas the requested price was not always retained, the metric should consider the final price.
- Regarding in which year to count the award, rejection, or write-off, the project partners considered the tradition by their finance departments of counting awards in the year that the award was made. However, a distinction was made by the project partners between accounting rules preferred by the finance departments, and the objective of Snowball Metrics to indicate performance against strategic objectives and peers. The decision was taken to base the recipe on year of application, to ensure that the success or failure of the application counts against the application itself, and also to ensure a value of the metric between 0% and 100%.

### 1.6.2 Tailoring Snowball Metrics to engage with researchers at Queen's University Belfast changed behaviour and improved research performance[27]

The research management team at Queen's University Belfast noticed that, despite being ahead of their peers on growth in Scholarly Output ❄, they were falling behind in citations which account for a large proportion of the scores used in global rankings. The university launched a collaboration with Elsevier to gain insights into how it compares to its peers; this collaboration provided evidence that helped change the culture and conversations around research metrics that support a new publication impact strategy and research ambition.

Queen's University Belfast looked at the citation impact and quantity of publications, as well as the proportion of international collaborations. Analysis showed that the lag in citation growth compared to their selected peers might be due to the journals its scientists were publishing in, with a relatively low proportion in the highest citation impact journals (as measured by Source Normalized Impact per Paper (SNIP) values), and also highlighted that while international collaboration was generally higher than peers, it was mostly with European authors, with relatively few collaborations with institutions in North America.

This information was used by the research management team to engage a growing number of researchers, which has led to a culture change across the university; the conversation has moved from indifference and scepticism amongst some researchers to one where it is acknowledged that there is a role for citations, and that there are important ways of using data to understand performance in relation to peers. The Queen's University Belfast team developed a publication impact strategy to help researchers target the higher impact journals, shifting the focus from quantity to quality. The university has subsequently noted higher levels of publications in top-tier journals, continued development of international networks and collaborations, and a higher Field-Weighted Citation Impact ❄ compared to many of its peer institutions. The gap in citation values between Queen's University Belfast and its benchmark peer group has narrowed considerably with citation growth levels significantly outstripping peer institutions in recent years.

### 1.6.3 International collaboration at Washington State University[28]

Washington State University studied the publication and funding record of their faculty and noted a substantial amount of papers involving international co-authors. Further investigation revealed very little internationally earmarked extramural funding (approximately 2%), despite about 30 percent of peer-reviewed publications involving international co-authors. These observations, combined with national data indicating that US research output is going down and losing global competitiveness, highlighted that information was not easily available to either university administrators or the researchers themselves.

Washington State University used Elsevier's SciVal[29], which incorporates Snowball Metrics, to analyze and help answer: who is funding these collaborations?; can we deepen these relationships?; and are we building the right strategic partnerships in the international arena? This work engendered a National Science Foundation-funded project to determine which metrics are helpful in determining the impact of strategic international collaborations. The university is developing a blueprint for institutions to use data that they are already mining and metrics that can help in decision-making and resource allocation.

27 Goodchild van Hilten. 2017. How personalized metrics can change behavior and improve research performance.
https://www.elsevier.com/connect/how-personalized-metrics-can-change-behavior-and-improve-research-performance

28 Arasu. 2014. A role for the library in awakening to the power and potential of institutional metrics for research.
https://libraryconnect.elsevier.com/articles/role-library-awakening-power-and-potential-institutional-metrics-research
29 https://www.elsevier.com/solutions/scival

### 1.6.4 Consolidating research-intensive medical schools[30]

The challenge of merging and streamlining five independent medical schools resulted from the mandate from the UK government that research-intensive medical schools should be consolidated to achieve efficiency in clinical services. There was an immediate need at Imperial College London to apply business principles to the academic research enterprise, and develop an evidence-based decision-making model that was agreed on, and supported by, the faculty; this in turn meant that the perceived threat to academic freedom, as well as the belief that the immeasurable could not be quantified, needed to be overcome. A bottom-up approach was the only way to start to overcome these barriers with measures developed and adopted by academics themselves. The outcome was that the Faculty of Medicine released an unproductive overhead, invested in new staff and quickly climbed to be the strongest UK medical school, as measured by any input or output research measure.

As the new medical faculty coalesced, Imperial College London began to monitor metrics such as Success Rate ✿ in applications for grants, and could prove that using an evidence-base to inform a strategic approach to funding applications had a huge positive effect on success rates. The evidence-based model also supported the detailed tracking of the loss of market share to competitors within an increasing base of funding in a particular area, and reverse that situation within 18 months.

### 1.7 Summary

1. **The aim of Snowball Metrics is to become the international standard that enables research-intensive universities to understand their strengths and weaknesses, so that they can build and monitor effective strategies by ensuring the comparison of apples with apples.** Snowball Metrics should be used together with qualitative input, such as peer review and expert opinion, to provide the most complete set of evidence into decisions being made.

2. **Snowball Metrics provide the opportunity for institutions to approach evaluation from their own perspective, and to think about how to best answer their questions and needs, rather than reacting to the needs of others.** The metric recipes are owned by research-intensive institutions, not by funders and agencies. Snowball Metrics are consequently based on all data sources available - institutional data sources, third party sources, and commercially available sources – and do not depend on a particular data source or supplier.

3. **The output of Snowball Metrics is a set of mutually agreed and tested methodologies: "recipes". These recipes are available free-of-charge and can be used by anyone for their own purposes and, if applicable, under their own business models.**

---

30 Green. 2013. Evidence-based decision making in academic research: the "Snowball" effect.
Academic Executive Brief 3: 12-14. https://academicexecutives.elsevier.com/articles/evidence-based-decision-making-academic-research-snowball-effect

# 2. Overview and denominators

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

- Display of Snowball Metrics
- Primary data sources
- Granularity of denominators
- Counting
- Citation counts
- Outputs included in the calculation of a Snowball Metric
- Currency conversion

Denominator definitions:

- Institution
- Discipline
- Researcher
- FTE (full-time equivalent) count
- Research student
- Funder-type
- Funding-type
- Time period
- Full-time or part-time research students
- Home or overseas research students
- Gender

CERIFication[31] of Snowball Metrics
Expression of Snowball Metrics in CASRAI data dictionary[32]

---

31 CERIF: Common European Research Information Framework;
   http://www.eurocris.org/Index.php?page=featuresCERIF&t=1

32 http://dictionary.casrai.org/Main_Page

## 2.1 Display of Snowball Metrics

A Snowball Metric is one which has been defined and agreed by research-focused universities as being useful in supporting strategic planning by enabling benchmarking between institutions. 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 organization, under their own business models if applicable.

A Snowball Metric is indicated by the use of this symbol ❄ placed after the name of the metric. This snowflake symbol can be downloaded from the Snowball Metrics website[33].

## 2.2 Primary data sources, and implications for benchmarking

The primary data sources listed are those that could be used to generate Snowball Metrics. Snowball Metrics recipes can be used regardless of the specific data sources available within a particular organization; for example, Scholarly Output ❄ could be generated using data from an institutional output repository or Current Research Information System (CRIS system), Scopus[34], Web of Science[35], or Google Scholar[36]. It is, however, important to have consistency in data sources when benchmarking between institutions to ensure that the comparisons are meaningful: it could 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, because differences could be caused by distinct database coverage, as well as performance.

For the Output Snowball Metrics, Institutional Output Repositories and CRIS systems[37] are referred to. These include Elsevier's Pure[38], Digital Science's Symplectic Elements[39], Clarivate Analytics' Converis[40], ePrints[41], and dSpace[42].

---

33 twww.snowballmetrics.com/metrics
34 https://www.elsevier.com/solutions/scopus
35 https://clarivate.com/products/web-of-science/
36 https://scholar.google.co.uk/
37 CRIS system: Current Research Information System
38 https://www.elsevier.com/solutions/pure
39 http://symplectic.co.uk/products/elements/
40 http://converis.thomsonreuters.com/
41 www.eprints.org
42 www.dspace.org

## 2.3 Granularity of denominators

Some metrics, when calculated for the more granular denominators, and especially for smaller institutions, will be based on few data points. These small data sets cause fluctuation in the metric between time periods, and may make the metrics less reliable for benchmarking.

It is recommended that information is shared along with the metric to highlight that a value is based on few data points when this is relevant, by populating the appropriate notes section of the Snowball Metrics Exchange service, or equivalent. This approach provides the option of reaching a high level of granularity for the metrics, while also providing the information for recipients of Snowball Metrics to determine how much weight to place on metrics calculated for granular denominators in benchmarking exercises.

### 2.3.1 UK application
Several UK applications of denominators are based on HESA data. It is recommended to apply the HESA rounding strategy of not calculating a metric from fewer than 7 data points. The following HESA rounding details are also noted:
- All numbers are rounded to the nearest multiple of 5.
- Any number lower than 2.5 is rounded to 0.
- Halves are always rounded upwards (e.g. 2.5 is rounded to 5).
- Percentages based on fewer than 22.5 individuals are suppressed.
- Averages based on 7 or fewer individuals are suppressed.
- The above requirements apply to headcounts and Full Time Equivalent (FTE) data.
- Financial data are not rounded.

## 2.4 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 D1, and C is a member of a separate disciplinary denominator D2:
- 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 D1, and 1 publication will also be credited to D2, when reading the metric out at these denominators. Fractional counting would credit both D1 and D2 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 D1 and D2 and so counted twice at the disciplinary denominator level.
- The only exception is Field-Weighted Citation Impact ❄ for which an output that is part of more than one subject field is counted fractionally in each, so that a single output does not exert too much weight on the metric value.

## 2.5 Citation counts

Some Snowball Metrics depend on counts of citations. These citation counts are typically the total number of citations received since publication up to the date of the current data extract.

The only exception is Field-Weighted Citation Impact ❄, which applies a current-plus-3-year citation window; for example, for an item published in October 2007, citations that are received in the remainder of 2007 until the end of December 2010 will be counted.

## 2.6 Outputs included in the calculation of a Snowball Metric

Every output in a data set would ideally be associated with the information needed for it to be included in the calculation of every Snowball Metric. In practice this is probably not the case; outputs in institutional repositories do not always have associated counts of citations or affiliation information, and outputs are not always part of serials that have journal metrics values, for example. All outputs that have the information needed to generate a Snowball Metric are included in the calculation, and outputs that lack the necessary information are excluded.

## 2.7 Currency conversion

The international benchmarking of financial recipes such as Applications Volume ❄, Awards Volume ❄ and Income Volume ❄ depend on an exchange rate mechanism. Federal reserve data will be used, and the exchange rates applied will be:
- The historical conversion rate for complete past years.
- The conversion rate at the time of the snapshot for the current year.

## 2.8 Institution denominator

An institution is defined as the sum of all units of an organization included in or with an organization's consolidated annual financial statements, and headed by one president, chancellor, or equivalent.

The *physical location of a campus* does not define whether something is or is not part of a university; overseas campuses are included, because they are recorded within the data elements of the institutional systems.

This definition could include:
- Agricultural experiment stations.
- Branch/international campuses.
- Medical schools, teaching hospitals or clinics.
- Research centers and facilities.
- University non-profit foundation (in US terminology, a "501(c)3 foundation"[43]).
- National research centers managed by the university.

This definition excludes:
- Any campus headed by a separate campus-level president, chancellor, or equivalent.
- Spin-out companies; for example, Google is not included within Stanford University, despite being founded there.
- Other organizations or institutions with which an organization has an affiliation or relationship, but which are not components of the organization nor included in or with its financial statements.

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.

Snowball Metrics support institutional decision making, and therefore data are viewed from an institutional standpoint. 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.

### 2.8.1 US application
This definition excludes federally funded research and development centers (FFRDCs[44]).

### 2.8.2 Future opportunity
Individual institutions could be aggregated on an as-needed basis within a tool, but the aggregations should not be considered unique institutional entities. For example:
- All of the campuses headed by a separate campus-level president, chancellor, or equivalent, such as the 3 campuses of the University of Illinois.
- State system institutions, such as the 10 institutions of the University of California system.
- Academic consortia, such as the Association of American Universities[45], the Association of Public and Land-grant Universities[46], the Council of Independent Colleges[47], the Ivy League universities[48], and so on.

---

43 https://www.irs.gov/charities-non-profits/charitable-organizations/exemption-requirements-section-501-c-3-organizations

44 http://www.nsf.gov/statistics/ffrdc/
45 https://www.aau.edu/
46 http://www.aplu.org/
47 https://www.cic.edu/Pages/default.aspx
48 https://en.wikipedia.org/wiki/Ivy_League

## 2.9 Discipline denominator

The discipline denominator enables benchmarking between institutions at a more granular level than that of the entire institution. A meaningful discipline-level denominator has the following characteristics:
- It is a structure that has the same meaning at all institutions.
- It draws on data that are readily available to an institution.
- It uses information that is reasonably current.

### 2.9.1 UK application
There are 2 options for the discipline denominator, and that which is relevant is specified in the "UK application" section of the particular recipe. They are both based on classifications produced by HESA[49].

#### 2.9.1.1 HESA cost centres
The HESA cost centre is a grouping of student, staff and finance records that is used as a framework for all institutions throughout the UK to return data annually to HESA. Cost centres do not reflect an institution's own organizational structure or strategic priorities, unlike departments and the Units of Assessment used by the Research Excellence Framework[50] exercise, making them ideal to support benchmarking between institutions.

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[51] is used to create this denominator for Snowball Metrics.

Students are attributed to all of the HESA cost centres of the academic staff that supervise and/or teach them.

#### 2.9.1.2 HESA JACS[52] codes
Joint Academic Coding System (JACS) codes are a way of classifying academic subjects and modules. JACS codes are maintained jointly by HESA and the Universities and Colleges Admissions Service[53] (UCAS). They are used at the 2-digit principal subject level.

JACS codes are due to be replaced by the Higher Education Coding System (HECoS) and the Common Aggregation Hierarchy (CAH) in the near future[54], in which case the new equivalent field should be used.

### 2.9.2 US application
The US application of the discipline denominator is based on the National Science Foundation's Higher Education Research and Development survey (HERD) classification[55].

### 2.9.3 UK-US discipline denominator mapping
The national funder-type denominators for the UK and the US may not be relevant for benchmarking outside these countries. The shared classification described in this section is based on a mapping of the national denominators, and can be used to benchmark between UK and US institutions.

The shared classification is the most simple solution that enables benchmarking between UK and US institutions, based on their national discipline denominators. It aims to conserve the information held within each national denominator, while requiring a minimal amount of effort that is considered feasible to achieve.

The HERD classification is more granular than the HESA classification, in that it has 3 levels rather than 1. Figure 4 shows a partial illustration of the mapping.

**Figure 4: Example of discipline mapping between the HESA and HERD classifications**
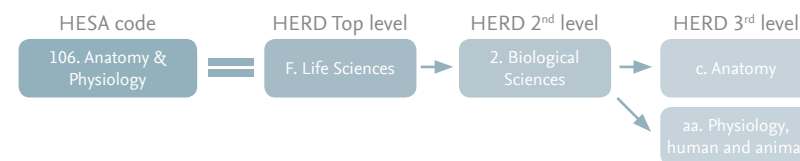


The complete HESA-HERD discipline mapping is too large to be reproduced in the recipe book, and so is available via the metrics tab of the Snowball Metrics website[56].

49 www.hesa.co.uk

50 www.ref.ac.uk/

51 https://www.hesa.ac.uk/files/staff_2011-12_c11026.pdf page 49.

52 www.hesa.co.uk/support/documentation/jacs

53 https://www.ucas.com/

54 JACS will be replaced by the Higher Education Costing System (HeCoS) and the Common Aggregation Hierarchy (CAH) for the 2019/20 academic year: www.hesa.ac.uk/innocation/hecos

55 www.nsf.gov/statistics/srvyherd/

56 www.snowballmetrics.com/metrics/

## 2.9.3.1 *Future opportunities*

A new more sophisticated international discipline denominator could provide some enhancements to the current HESA-HERD mapping:

- All HESA classes could be mapped to the shared classes automatically. The only severe conflict between the HESA and HERD classifications is that on the HERD side, nothing corresponds to the HESA class Veterinary Science: Veterinary Medicine objects are distributed over 3 classes (Agricultural Sciences, Biological Sciences and Medical Sciences), and more Veterinary fields are to be found under F2 (Life Sciences / Biological sciences: Pathology, human and animal; Pharmacology, human and animal; Physiology, human and animal). Comparable figures for Veterinary Sciences would require a reclassification of objects into an agreed international denominator.

- Some areas, such as Medicine, Engineering, and Humanities, seem to be more differentiated than others.

Such a mapping would require significantly more effort on both sides since a new scheme would need to be agreed, and there would also need to be agreement regarding the method by which objects would need to be reclassified. Reclassification could be done without software support (manually), or with partial or complete software support. If such a classification will ever be attempted, it is recommended that it is done once experience based on benchmarking with the above HESA-HERD mapping has been gained.

## 2.10 Researcher denominator

A researcher is any faculty or staff member who could act as the principal investigator of a funding application and who spends >0% time on research.

This definition includes all those working in research-focused universities who have time allocated to research of any kind, such as:
- Researchers who engage in "traditional" laboratory work.
- Clinicians who are doing even a small amount of research.
- Librarians and professional research staff, such as research associates who are performing research solely with internal or philanthropic funds.

This definition excludes trainees including undergraduate and graduate students, post-doctoral researchers, and staff or faculty with limited-term or temporary appointments such as visiting scholars.

Consideration has been given to the specification of a census date upon which a researcher must be in institutional systems for their outputs to be considered towards Snowball Metrics. However, the conclusion was that this is not practical and it is therefore not part of the researcher denominator. It is recommended that the date of the data on which this denominator is based is shared in the information exchanged with Snowball Metrics for benchmarking, for example in the Snowball Metrics Exchange service.

### 2.10.1 UK application
A researcher is any institutional employee whose contract of employment, as defined by the Higher Education Statistics Agency's (HESA[57]) Academic Employment Function field (ACEMPFUN), is either "2: Research-only" or "3: Teaching and research", and who is also not flagged in the HESA Research Assistant field (RESAST) as "1: Research assistant".

### 2.10.2 US application
This is drawn from the annual federal Time & Effort Reporting, directed by the US Office of Management and Budget[58].

---

57 www.hesa.ac.uk
58 http://www.whitehouse.gov/omb/

## 2.11 FTE (full-time equivalent) count denominator

FTE count indicates the extent of a researcher's workload that is focused on research.

FTE count is used to provide the option to normalize for different sizes of institutions, and disciplines within those institutions. The FTE normalization option is included for those Snowball Metrics for which a larger institution or discipline would generally be expected to do more of whatever is being measured, and it not included for those Snowball Metrics which are most likely to be affected by factors other than size, such as type of institution, disciplinary focus, strategy and mission statement. For example:

- Scholarly Output ❄ is very strongly related to size, and so the recipe includes FTE normalization.
- Sustainable Spin-Offs ❄ is only loosely related to the research activities and the researchers that conducted them, and so the recipe does not include FTE normalization.

### 2.11.1 UK application
The FTE count of those Researchers returned by institutions to HESA.

### 2.11.2 US application
This is drawn from the annual federal Time & Effort Reporting, directed by the US Office of Management and Budget[59].

## 2.12 Research student denominator

A research student is any student studying for either a doctoral award or a masters degree by research, having achieved a first degree as a condition of entry. A research-based higher degree is a postgraduate programme comprising a research component (including a requirement to produce original work) that is larger than any accompanying taught component when measured by student effort.

This definition excludes trainees including graduate students undertaking taught courses, post-doctoral researchers, and staff or faculty with limited-term or temporary appointments such as visiting scholars.

### 2.12.1 UK application
A research-based higher degree is a postgraduate programme comprising: 1 – a research component (including a requirement to produce original work) and, 2 – a research component that is larger than any accompanying taught component when measured by student effort. The arrangements for assuring and maintaining the academic standards and enhancing the quality of these programmes should be fully compliant with chapter B11 (Research degrees) of the Quality Assurance Agency for Higher Education's "UK Quality Code for Higher Education"[60].

Some specialist doctoral degrees, such as Doctor of Education (EdD) and Doctor of Clinical Psychology (DClinPsychol), include a research component but also include significant taught components and supervised practice. These degrees do not generally require the student to produce the same amount of original research as a PhD. Students registered for a specialist doctoral degree should only be included as research students if they satisfy both criteria.

A research student is based on the HESA COURSEAIM[61] field. This field describes the general qualification aim of the course and is intended to record the qualification that will be attained as a result of successful completion of studies.

---

59 http://www.whitehouse.gov/omb/

60 http://www.qaa.ac.uk/assuring-standards-and-quality/the-quality-code
61 https://www.hesa.ac.uk/index.php?option=com_studrec&task=show_file&mnl=13051&href=a^_^COURSEAIM.html

The following COURSEAIM codes will be used:
- D00 – doctorate degree that meets the criteria for a research-based higher degree
- D01 – new route PhD that meets the criteria for a research-based higher degree
- L00 – masters degree that meets the criteria for a research-based higher degree
- L80 – other postgraduate qualification at level L that meets the criteria for a research-based higher degree
- L99 – advanced supervised research at levels D or L with an unspecified qualification aim

The following COURSEAIM codes are excluded:
- L91 – visiting research students at levels D or L, with formal or informal credit
- D90 – advanced supervised research at level D for institutional credit
- L90 – advanced supervised research at level L for institutional credit

### 2.12.2 Future opportunity
It would be valuable to define post-doctoral researchers as a distinct denominator.

## 2.13 Funder-type denominator

This denominator is applied to:
- Applications Volume ❄: to the count, price, or amount applied for.
- Awards Volume ❄: to the count and value of awards.
- Income Volume ❄: to the income spent.
- Market Share ❄: to research income.

### 2.13.1 UK application

| Snowball Metrics denominator | Constituent HESA Funder Types | Further Breakdown for Snowball Metrics |
|---|---|---|
| Research Councils | Research Councils, Royal Society & British Academy | • AHRC<br>• BBSRC<br>• EPSRC<br>• ESRC<br>• MRC<br>• NERC<br>• STFC<br>• British Academy<br>• Royal Society |
| UK Charity | • UK-based Charity (QR Eligible for Charities Support)<br>• UK-based Charity (NOT QR eligible) | |
| UK Public Sector | UK central government bodies/ local authorities, health & hospital authorities | |
| UK Industry | UK industry, commerce & public corporations | |
| Overseas Industry | • EU industry, commerce & public corporations<br>• Non-EU industry, commerce & public corporations | |
| EU Government | EU government bodies | European Commission Other EU government bodies (in aggregate) |
| Other Overseas Sources | • EU-based charities (QR Eligible for Charities Support)<br>• Non-EU-based Charity (QR Eligible for Charity Support)<br>• EU other<br>• Non-EU other | |
| Other Sources | Other sources | |

## 2.13.2 US application

The US application of the funder-type denominator is based on the FundRef classification[62].

| Group | Group description | Abbrev. | Class | Class description |
|---|---|---|---|---|
| Government | All federal grant-making institutions and selected state or other (local) government agencies | govfed | Federal / national government | Governmental bodies on a national level (national governments and national governmental departments) |
| Government | | govnon | Government Non-federal | Governmental bodies on a non-national level (departments of states, cities, et cetera) |
| Private | all private foundations providing research funding; either corporate, family or other type of philanthropic giving | founda | Foundation | Every funding body named 'foundation' |
| Corporate | Corporations providing research grants and contracts | corpor | Corporate | (Large) cooperations, often only fund research 'on the side' |
| Other | International, non-governmental organizations (NGOs), other non-profits | academ | Academic | Universities, colleges – internal funding |
| Other | | assoc | Professional Associations and Societies | Every association or society of professionals in a field (including associations/societies of academic professionals) |
| Other | | intern | International | All other funding bodies, not residing in one specific country (without headquarters, or with multiple headquarters in several countries) |
| Other | | othern | Other Non-profit | All other funding bodies |

## 2.13.3 UK-US funder-type mapping

The national funder-type denominators for the UK and the US may not be relevant for benchmarking outside these countries. The following shared classification, that is based on the national denominators, can be used to benchmark between UK and US institutions.

| Shared classification | FundRef | HESA funder-types |
|---|---|---|
| Federal / National Government | govfed | Research Councils |
| Government Non-Federal | govnon | UK public sector |
| Foundation | founda | UK charity |
| Corporate | corpor | UK industry Overseas industry |
| Academic | academ | Other sources |
| Professional Associations and Societies | assoc | Other sources |
| International | intern | EU Government Other overseas sources |
| Other sources | othern | Other sources |

---

62 https://www.crossref.org/services/funder-registry/

## 2.14 Funding-type denominator

Funding-type is a denominator for the Research Student Funding ❄ recipe.

### 2.14.1 UK application

This denominator is derived from the HESA "major source of tuition fees (MSTUFE)[63]" classification.

| Snowball Metrics | Constituent HESA Funding-Types |
|---|---|
| No award or financial backing | No award or financial backing |
| Institutional waiver of support costs | Institutional waiver of support costs, including university scholarships or awards of current institution (from HESA notes) |
| UK & Islands Government Bodies | • Local Government - Channel Islands & Isle of Man/Scottish FE Bursaries<br>• Departments of Health/NHS/Social Care<br>• Departments of Social Services<br>• Department for Business, Energy and Industrial Strategy (BEIS; formerly the Department of Business, Innovation and Skills)<br>• Other HM government departments/public bodies<br>• Scholarship of HM forces<br>• Scottish Enterprise Network/Highlands & Islands Enterprise/ Local Enterprise Companies (LECs)<br>• LEA training grants scheme<br>• Department of Agriculture & Rural Development for Northern Ireland (DARDNI)<br>• Scottish Local Authority - discretionary award<br>• Overseas student award from HM government/British Council<br>• Department for International Development |
| Research Councils & British Academy | • British Academy<br>• Biotechnology & Biological Sciences Research Council (BBSRC)<br>• Medical Research Council (MRC)<br>• Natural Environment Research Council (NERC)<br>• Engineering & Physical Sciences Research Council (EPSRC)<br>• Economic & Social Research Council (ESRC)<br>• Arts & Humanities Research Council (AHRC)<br>• Science & Technology Facilities Council (STFC)<br>• Research council - not specified |

| Snowball Metrics | Constituent HESA Funding-Types |
|---|---|
| Charitable Foundations | • Cancer Research UK<br>• Wellcome Trust<br>• Other Association of Medical Research Charities (AMRC) charity<br>• Other charitable foundation |
| Overseas Sources | • International agency<br>• EU Commission (EC)<br>• Overseas government<br>• Overseas institution<br>• Other overseas funding<br>• Other overseas - repayable loan |
| Industry & Commerce (UK & overseas) | • Overseas industry or commerce<br>• UK industry/commerce |
| Other Sources | • Fee waiver under government unemployed students scheme<br>• Student's employer<br>• Other<br>• No fees<br>• Not known |

---

63 https://www.hesa.ac.uk/collection/c16051/e/mstufee

## 2.15 Time period denominator

**2.15.1 Calendar year**
A calendar year runs from 1 January to the earliest 31 December thereafter.

**2.15.2 Financial year**
UK application: 1 August to 31 July of the following year.

**2.15.3 Quarter**
A 3-month period, applied to Applications Volume ❄ and Awards Volume ❄.

*2.15.3.1 UK application*
- 1 August to the earliest 31 October thereafter
- 1 November to the earliest 31 January thereafter
- 1 February to the earliest 30 April thereafter
- 1 May to the earliest 31 July thereafter

## 2.16 Full-time or part-time research students denominator

This is the method by which a student is being taught their course.

**2.16.1 UK application**
This denominator is drawn from the HESA MODE[64] field.

---

## 2.17 Home or overseas research students denominator

This provides the option to look at the student population at a more granular level, separating home and overseas students.

### 2.17.1 UK application
The HESA fee eligibility (FEELIG)[65] field is used to distinguish between students who are eligible to pay home fees from those who are not.

## 2.18 Gender denominator

The gender denominator is used to monitor equal opportunities issues.

### 2.18.1 UK application
The HESA SEXID[66] field is used. The valid entries for Snowball Metrics are:
- 1 – male.
- 2 – female.

Code 3 – "Other" will not be used in Snowball Metrics
- Equality Challenge Unit (ECU)[67] recommends the use of the terms "other" and "prefer not to say" for people who associate with the terms intersex, androgyne, intergender, ambigender, gender fluid, polygender and genderqueer. HESA does not include a "prefer not to say" option.
- There are very few research students classified as "other" across the sector, and the numbers are considered too low for robust benchmarking.

---

65 https://www.hesa.ac.uk/collection/c14051/a/feeelig/

66 https://www.hesa.ac.uk/index.php?option=com_studrec&task=show_file&mnl=13051&href=a^_^SEXID.html

67 http://www.ecu.ac.uk/

## 2.19 CERIFication of Snowball Metrics

The Common European Research Information Format[68] (CERIF), is a global data standard developed by euroCRIS[69], a not-for-profit organization that is dedicated to the development and inter-operability of Current Research Information Systems (CRIS systems). euroCRIS continues to partner with Snowball Metrics to express recipes in CERIF.

The CERIFied recipes have been used in the implementation of the Snowball Metrics Exchange service (section 1.4.1). This work in partnership between Elsevier and euroCRIS highlighted some necessary enhancements to CERIF xml to improve its usability for the purposes of calculating metrics, which are incorporated into CERIF by euroCRIS.

The complete CERIF xml code for Snowball Metrics, as prepared by euroCRIS, is available for download and use from the Snowball Metrics website[70]. The accuracy of CERIFied Snowball Metrics is the responsibility of euroCRIS.

The following principles have been applied in the CERIFication of Snowball Metrics[71]:

- The CERIFication is performed by, and approved by, the euroCRIS CERIF Task Group.
- The CERIFication is carried out on the generic metric definition to ensure global relevance, rather than on the national applications.
- The principle of Snowball Metrics being system-agnostic is followed in that the data source is described, but not the methodology that a specific system uses to generate the metric value.
- The final calculated Snowball Metric will be exchanged between systems, rather than the separate components needed to generate the value. For example, when normalizing by FTE count, the normalized value is exchanged, and not the metric plus the separate FTE count from which the recipient would need then to complete the normalization themselves.

The hierarchy of the original set of 10 Snowball Metrics is shown in Figure 5.

The generic translation of the Snowball Metrics recipes to CERIF is shown in Figure 6.

Figure 5: The hierarchy of the original set of 10 Snowball Metrics



---

68 http://www.eurocris.org/Index.php?page=featuresCERIF&t=1
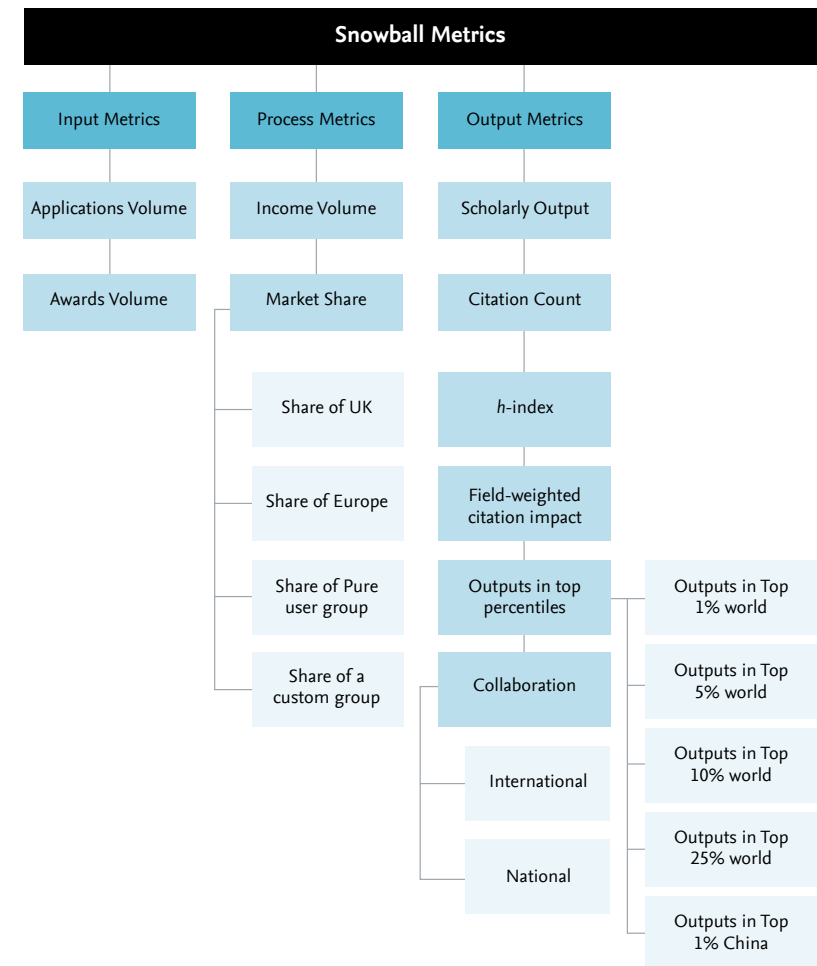
69 www.eurocris.org

70 www.snowballmetrics.com/metrics/

71 Clements, Jorg, Lingjaerde, Chudlarsky, Colledge, 2014. The application of the CERIF data format to Snowball Metrics. Procedia Computer Science 33: 297-300. http://www.sciencedirect.com/science/article/pii/S1877050914008370

**Figure 6: Generic translation of Snowball Metrics recipes to CERIF**
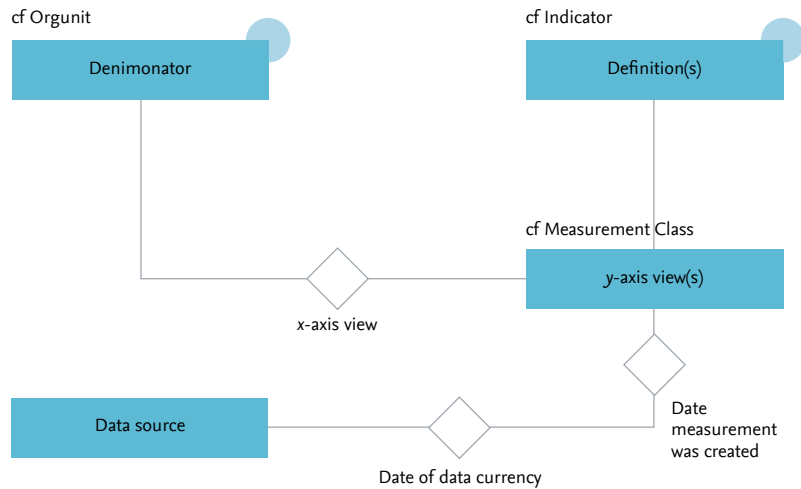


## 2.20 Expression of Snowball Metrics in CASRAI data dictionary

The Consortia Advancing Standards in Research Administration Information[72] (CASRAI), is an international not-for-profit membership initiative led by research institutions and their partners. It is dedicated to adapting the principles and best practices of open standards and data governance to develop "standard information agreements" that serve as bridges between research information users.

Through a project with CASRAI made possible by a donation from Elsevier, the Snowball Metrics project partners were able to engage with a larger group of universities, and also receive input into the metric recipes from funding bodies such as The Wellcome Trust[73] and Medical Research Council[74], other suppliers including Digital Science[75] and Thomson Reuters[76] as well as organizations such as RAND Europe[77]. The accuracy and completeness of standard information agreements for Snowball Metrics is the responsibility of CASRAI.

72 http://casrai.org/

73 https://wellcome.ac.uk/

74 https://www.mrc.ac.uk/

75 https://www.digital-science.com/

76 https://www.thomsonreuters.com/en.html

77 https://www.rand.org/randeurope.html

# 3. Input Metrics

This section details the methodologies for the following Snowball Metrics:

Research Metrics:
1. Applications Volume
2. Awards Volume
3. Success Rate

Enterprise Activities / Economic Development Metrics:
4. Academic-Industry Leverage
5. Business Consultancy Activities

Post-Graduate Research Metrics:
6. Research Student Funding

## 3.1 Applications Volume ❄

VOLUME OF RESEARCH GRANT APPLICATIONS SUBMITTED TO EXTERNAL FUNDING BODIES

ENDORSED BY: UNITED KINGDOM.

### 3.1.1 Metric definition
Applications Volume ❄ calculates the number and price[78], or amount applied for[79], of research grant applications that are submitted to external funding bodies.

It answers the questions of:
- Whether the institution is applying for more or less research funding than in the past.
- The funder type profile of grant applications.



(a) Count of applications
(b) Price of applications, or amount applied for

(a) Time period

(a) Count of applications per FTE
(b) Price of applications, or amount applied for, per FTE

(a) Time period
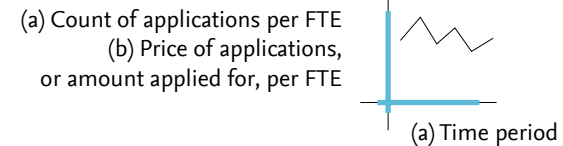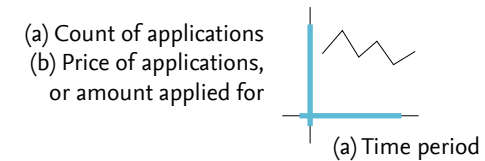
The complete Common European Research Information Format (CERIF)[80] xml code for this metric, prepared by euroCRIS, is available for download and use via the Snowball Metrics website[81].

---

78 "Price" is the phrase typically used in the United Kingdom.

79 "Amount applied for" is the phrase typically used in Australia / New Zealand.

80 http://www.eurocris.org/Index.php?page=featuresCERIF&t=1

81 www.snowballmetrics.com

### 3.1.2 Details

The price, or amount applied for, 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 avoids double-counting of the same applications by excluding prior submissions in a multi-stage application process such as outlines and expressions of interest. For example:

- A £1m application is submitted as an expression of interest. At this stage, it should be included in Applications Volume ❄.
- If the application is declined, then the expression of interest is considered to be the application, and is counted within Applications Volume ❄.
- If the application is reviewed favorably and invited to proceed to the full submission stage, then the full application is considered to be the application, and replaces the expression of interest. It is not the intention for the £1m application to be considered as both an expression of interest and as a full application, but only counted once as the most recent.

Supplements should be treated as new applications. For example:

- Consider an application for a total of £1m, to start in financial year 2007/2008.
- If there is a single application, this will be recorded as £1m in financial year 2007/2008.
- If there are 5 annual applications, each of £200k, then a new application of £200k is recorded for each of the financial years 2007/2008, 2008/2009, and so on to 2011/2012.

The date used is the date on which the application is submitted to the funding body.

### 3.1.3 Primary data source

- Institutional research grant application system or Current Research Information System (CRIS system)

### 3.1.4 UK application

Denominators derived from institutional data:

- 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 at the institution.
- Funder-type
- Institution

Time period:

- Financial year
- Quarter of financial year

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[82]. 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, and EU Partner elements.
- Any research that could not be classified as "Research" under HESA's accounting conventions for the finance record (Financial Reporting Standard 102)[83].

Currency: British pounds (GBP).

Awards considered in Applications Volume ❄ should be in line with the guidance provided by HESA in their table "Research grants and contracts – breakdown by source of income and HESA cost centre"[84].

### 3.1.5 US application

The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

The FundRef-based classification is to be used as a counterpart to the HESA funder-type mapping. See also the UK-US funder-type mapping in section 2.13.3, that enables benchmarking between UK and US institutions based on national funder-type data.

---

82 http://www.hesa.ac.uk/index.php?option=com_content&task=view&id=1145&Itemid=233

83 https://www.hesa.ac.uk/collection/c15031/coverage#conventions

84 https://www.hesa.ac.uk/collection/c15031/table_5

### 3.1.6 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 a subset of cases, institutions also capture co-investigators in their grant application systems; the principal investigator approach was agreed since it is inclusive and ensures that all institutions can use the same methodology. It is an interesting opportunity for the future to consider a mapping according to co-investigators as well as to the principal investigator.

A denominator reflecting themes and subject focus of the competitive funding applications would be highly valued, especially if the same thematic denominator could be applied not only to Input, but also to Process, and Output and Outcome, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

A critical mass of national funding bodies might be considered a source of data for this metric. However, particular attention would need to be paid to the consistency and compatibility of approach in recording the data at each participating funding body to ensure the resultant information is comparable and meaningful.

- For example, funders often allocate the full value of an application to the lead institution within their databases:
  - If Institution 1 was the leading institution on a GBP 10 million application of which Institutions 2, 3 and 4 were also receiving equal portions, then Institutions 2, 3 and 4 would each have GBP 2.5 million too little in their Applications Volume ❄ and Institution 1would show GBP 7.5 million too much.
  - There are large applications that could significantly affect the ability to benchmark using Awards Volume ❄, if they are allocated entirely to the lead institution.

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

## 3.2 Awards Volume ❄

VOLUME OF RESEARCH AWARDS GRANTED AND AVAILABLE TO BE SPENT

ENDORSED BY: UNITED KINGDOM.

### 3.2.1 Metric definition

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

It answers the questions of:
- Whether the institution is winning more or less in research grants and contracts than in the past.
- The funder type profile of awarded grants.



(a) Count of awards
(b) Value of awards

(a) Time period

(a) Count of awards per FTE
(b) Value of awards per FTE

(a) Time period
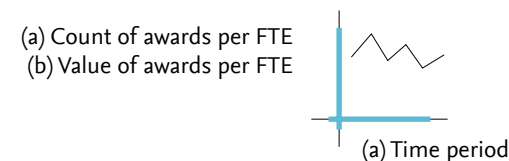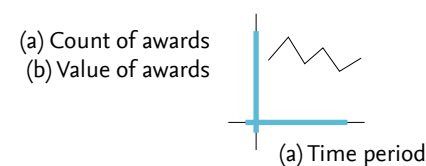
The complete Common European Research Information Format (CERIF)[85] xml code for this metric, prepared by euroCRIS, is available for download and use via the Snowball Metrics website[86].

### 3.2.2 Details

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

---

85 http://www.eurocris.org/Index.php?page=featuresCERIF&t=1
86 www.snowballmetrics.com

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 (i.e. zero-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 financial year 2007/2008.
- If this award is increased by £0.5m in financial year 2010/2011, an award of £0.5m, not £1.5m, is recorded in financial year 2010/2011.
- If the award is then reduced by £0.2m in financial year 2011/2012, an award of -£0.2m, not £1.3m, is recorded in financial year 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 on which the award is entered in the institutional grants system.

### 3.2.3 Primary data source
- Institutional grants system or Current Research Information System (CRIS system).

### 3.2.4 UK application
Denominators derived from institutional data:
- 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 at the institution.
- Funder-type
- Institution

Time period:
- Financial year
- Quarter of financial year

Awards considered reflect activities where the resultant spend would be returned as research grants and contracts income in the HESA financial return[87].

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 / Partnerships, and EU Partner elements.
- Any research that could not be classified as "Research" under HESA's accounting conventions for the finance record (Financial Reporting Standard 102)[88].

Currency: British pounds (GBP).

Awards considered in Awards Volume ✿ should be in line with the guidance provided by HESA in their table "Research grants and contracts – breakdown by source of income and HESA cost centre"[89].

### 3.2.5 US application
The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in the section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

The FundRef-based classification is to be used as a counterpart to the HESA funder-type mapping. See also the UK-US funder-type mapping in section 2.13.3, that enables benchmarking between UK and US institutions based on national funder-type data.

### 3.2.6 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 a subset of cases, institutions also capture co-investigators in their grant application systems; the principal investigator approach was agreed since it is inclusive and ensures that all institutions can use the same methodology. It is an interesting opportunity for the future to consider a mapping according to co-investigators as well as to the principal investigator.

87 http://www.hesa.ac.uk/index.php?option=com_content&task=view&id=1145&Itemid=233

88 https://www.hesa.ac.uk/collection/c15031/coverage#conventions

89 https://www.hesa.ac.uk/collection/c15031/table_5

A denominator reflecting themes and subject focus of awards granted would be highly valued, especially if the same thematic denominator could be applied not only to Input, but also to Process, and Output and Outcome, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

A critical mass of national funding bodies might be considered a source of data for this metric. However, particular attention would need to be paid to the consistency and compatibility of approach in recording the data at each participating funding body to ensure the resultant information is comparable and meaningful.

- For example, funders often allocate the full value of an award to the lead institution within their databases:
  - If Institution 1 was the leading institution on a GBP 10 million award of which Institutions 2, 3 and 4 were also receiving equal portions, then Institutions 2, 3 and 4 would each have GBP 2.5 million too little in their Awards Volume ❄ and Institution 1 would show GBP 7.5 million too much.
  - There are large awards that could significantly affect the ability to benchmark using Awards Volume ❄ , if they are allocated entirely to the lead institution.

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

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. It may also result in values from previous periods being amended, which can be difficult to explain to recipients of this information.

## 3.3 Success Rate ❄

PROPORTION OF RESEARCH GRANT APPLICATIONS THAT HAVE BEEN SUCCESSFUL

ENDORSED BY: UNITED KINGDOM.

### 3.3.1 Metric definition
Success Rate ❄ calculates the proportion of research grant applications, submitted to external funding bodies, which have been successful.

It answers the questions of:
- How successful is the institution in winning research grants and how is this changing over time.
- Whether the institution is more or less successful at winning smaller or larger grants and whether there is a trend.

The Success Rate ❄ calculated by this method will change over time:
- Success Rate by count is calculated according to whether submitted applications have been awarded or rejected, or whether a decision is pending.
- Success Rate by value is calculated according to the proportion of the total requested price associated with awarded or rejected applications or whether a decision is pending.

This method was of concern to colleagues in finance who are used to year end numbers being absolute. This method will only provide absolute values when decisions have been received on all applications submitted in a given year. The method of tying Success Rate to application year rather than to award year has been selected to avoid the problem of values being greater than 100%.



(a) Successful applications by count
(b) % pending applications by count
(c) % rejected applications by count

(a) Application year



(a) Successful applications by value
(b) % pending applications by value
(c) % rejected applications by value
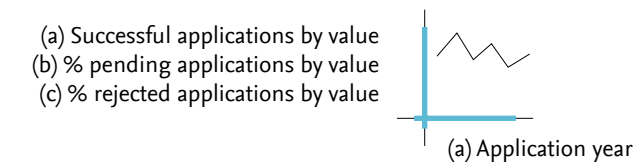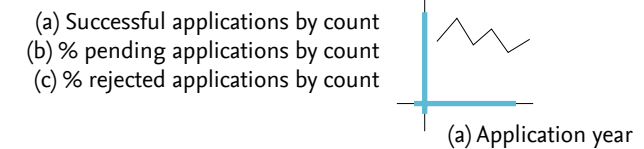
(a) Application year

The 3 variants on the y-axis for each chart above could be displayed in a composite chart, as suggested below.

**Success Rates of applications to Research Councils from cost centre of chemistry**



*Chart: x-axis labelled 2013/14, 2014/15, 2015/16, 2016/17 for University A, University B, University C; "Financial year of submission". y-axis "Research Grant Applications – % of Count" 0–100. Legend: Awarded, Pending, Rejected.*
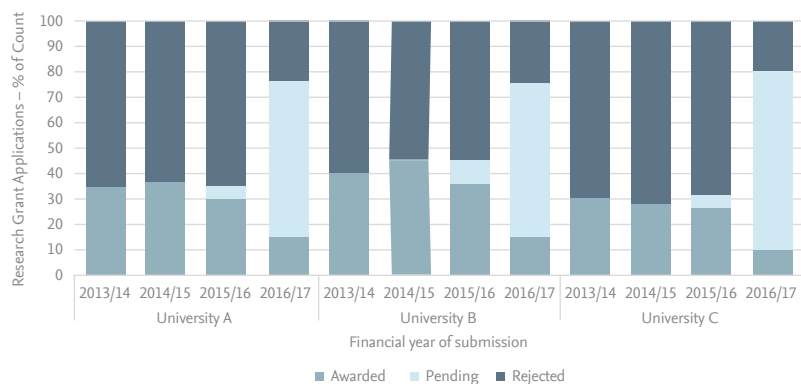
### 3.3.2 Details

The set of awards considered in this metric is different to the set of awards counted in the Awards Volume ❄ metric.

The price used should be the most up-to-date available. For example, if the price is revised by the funder or in negotiations with research partners, this new price should be used as the basis for this calculation, by adjusting the price requested to match the price awarded.

Awards that are not tied to an application do not count towards Success Rate.

If an application is added retrospectively into institutional systems ("dummy application") in order to improve record keeping, then this application and its outcome will be counted.

The year of success is the date of submission to the funder. If an award was applied for in 2012, and it was awarded in 2013, the Success Rate will be tied back to the original 2012 application year.

The following awards are not considered in Success Rate ❄:
* Donations that were not applied for.
* Awards that were transferred in from other institutions when a researcher moves (note that this *would* be included in Awards Volume ❄).
* Open access grants such as those provided by Research Councils UK[90] from time to time, and similar.
* Any research that could not be classified as "Research" under HESA's accounting conventions for the finance record (Financial Reporting Standard 102)[91].
* It is noted that not all systems enable these exclusions. It is expected that their volume is low and will not have a significant effect on the metric, and so this technically correct detail within the metric recipe is adopted. Please add a note into the Snowball Metrics Exchange service, when exchanging this metric, if donations have not been excluded.

Success Rate ❄ addresses new applications only, just like Applications Volume ❄. It avoids double counting of the same applications by excluding prior submissions in a multi-stage application process such as outlines and expressions of interest.
* In situations where Expressions of Interest etc. are not recorded in the Awards Management System, it is difficult to avoid double counting. It is expected that their volume is low and will not have a significant effect on the metric, and so this technically correct detail within the metric recipe is adopted. Please add a note into the Snowball Metrics Exchange service, when exchanging this metric, if prior submissions have not been excluded.
* Competitive renewals are considered to be new applications.

Success Rate ❄ allows for the amendment of the metric, and this is considered independent of Awards Volume ❄:
* Supplements are included in the calculation. If a supplement is applied for, then this counts as a new application in its own right and is assigned to the year in which it was applied for, not the year that the original award was applied for. For example, if an award was applied for in financial year 2011/2012 and awarded, and a supplement was subsequently received in financial year 2012/2013, then the supplement is considered in 2012/2013, and not 2011/2012).
* Underspend / returns are not considered, because the institution was awarded the money regardless of whether it was spent or not.
* No-cost extensions (i.e. zero-cost extensions) are not included.

90 http://www.rcuk.ac.uk/
91 https://www.hesa.ac.uk/collection/c15031/coverage#conventions

Write-off period: the definition assumes that an application has been declined where no decision has been entered into the system after 12 months. The reasons are:

- In order to be able to compare 'like with like' with external peers, the Snowball Metric needs an agreed write-off period. Institutions are still at liberty to use their own rules for internal views of the metric.
- It is important that there is a consistent time per funder, discipline and geography.

Note that the write off period does not apply if it is known that an award will be made but this has not yet been entered into the system, for example because it is in contract negotiation.

An example of Success Rate ❄ follows. Consider an institution that has made 7 applications in a given year:

| Application ID | Price | Successful? | Budget |
|---|---|---|---|
| 1 | 1,000,000 | Yes | 850,000 |
| 2 | 500,000 | Yes | 500,000 |
| 3 | 750,000 | No | N/A |
| 4 | 200,000 | No response 14 months after submission | N/A |
| 5 | 500,000 | No | N/A |
| 6 | 25,000 | Yes | 25,000 |
| 7 | 50,000 | No response 9 months after submission | N/A |
| **TOTAL** | **3,025,000** | | **1,375,000** |

- Success Rate by count
  - 42.9% (3/7) successful
  - 42.9% (3/7) rejected
  - 14.3% (1/7) pending
  - Note – application 7 is considered pending, and application 4 is considered rejected, according to the detail on the write-off period.
- Success Rate by value
  - 50.4% (1,525,000 / 3,025,000) successful
  - 47.9% (1,450,000 / 3,025,000) rejected
  - 1.7% (50,000 / 3,025,000) pending

### 3.3.3 Primary data source
Institutional research grant application system or Current Research Information System (CRIS system).

### 3.3.4 UK application
Denominators derived from institutional data:

- 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 at the institution.
- Funder-type
- Institution

Time period:
- Financial year
- Quarter of financial year

Currency: British pounds (GBP).

### 3.3.5 US application
The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

The FundRef-based classification is to be used as a counterpart to the HESA funder-type mapping. See also the UK-US funder-type mapping in section 2.13.3, that enables benchmarking between UK and US institutions based on national funder-type data.

*"It was great to be able to work with the Snowball Metrics Steering Group to help define the Success Rate metric – I was able to successfully incorporate the specific metric here at my own institution (Washington State University) and was pleased to be part of such a dedicated and knowledgeable team who have obviously put so much rigorous work into the recipes for Snowball Metrics."*
Derek A. Brown, Washington State University

### 3.3.6 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 a subset of cases, institutions also capture co-investigators in their grant application systems; the principal investigator approach was agreed since it is inclusive and ensures that all institutions can use the same methodology. It is an interesting opportunity for the future to consider a mapping according to co-investigators as well as to the principal investigator.

A denominator reflecting themes and subject focus of awards granted would be highly valued, especially if the same thematic denominator could be applied not only to Input, but also to Process, and Output and Outcome, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

## 3.4 Academic-Industry Leverage ❄

PRIVATE INVESTMENT LEVERAGED FROM PUBLIC SPONSORSHIP

ENDORSED BY: UNITED KINGDOM.

### 3.4.1 Metric definition

Academic-Industry Leverage ❄ calculates the total income for publicly sponsored research projects that are performed in collaboration with at least one other non-academic organization, and the percentage of funds from non-academic collaborators that this is used to leverage.

It answers the questions of:
- How much funding an institution is receiving to drive research and development through academic-industry partnerships.
- How effectively an institution is leveraging private investment in research and development from public funds.



(a) Total collaborative research income
(b) Non-academic contribution as a percentage of the amount of public funding

(a) Time period



(a) Total collaborative research income per FTE

(a) Time period
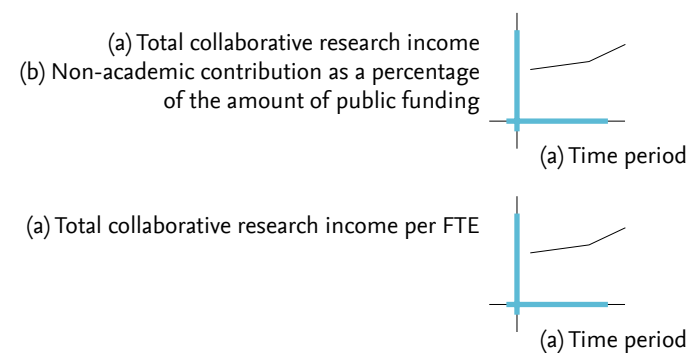
### 3.4.2 Details

Academic-Industry Leverage ❄ considers income associated with research projects that are publicly sponsored, and that are performed in collaboration with at least one other non-academic organization. A publicly sponsored research project is one which is funded by grant-in-aid from a Government or other public body. The collaboration should include material contribution, whether cash or "in kind", from at least one external non-academic collaborator.

The total income is the sum of income from both public funding, and from non-academic collaborators.

The non-academic contribution is the sum of cash and in kind contributions.

### 3.4.3 Primary data sources
- Institutional accounts system or Current Research Information System (CRIS system)
- Published annual accounts
- National statutory reports, such as those available from the Higher Education Statistics Agency[92] (HESA) in the UK

### 3.4.4 UK application

Denominator derived from institutional data: institution.

Time period: financial year.

Currency: British pounds (GBP).

---

## 3.5 Business Consultancy Activities ❄

VOLUME OF BUSINESS ENGAGEMENTS

ENDORSED BY: UNITED KINGDOM.

### 3.5.1 Metric definition

Business Consultancy Activities ❄ calculates the number and value of business engagements.

It answers the questions of:
- How much commercial income an institution is driving from consultancy.
- How effectively an institution is developing engagements with industry.



(a) Count of engagements
(b) Value of engagements

(a) Time period

(a) Count of engagements per FTE
(b) Value of engagements per FTE

(a) Time period
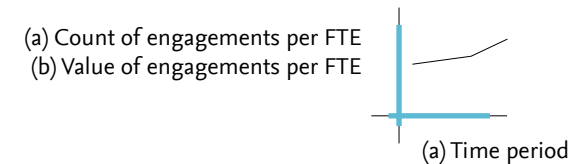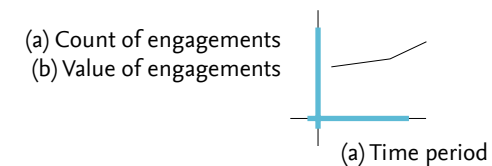
### 3.5.2 Details

Business consultancy is defined as the provision of expert advice and work that depends crucially on a high degree of intellectual input from the institution to the commercial or non-commercial client without the creation of new knowledge, even though the consultancy activities may involve a high degree of analysis, measurement and/or testing.

All consultancy activities where there is income to the institution should be considered, regardless of the contract-type of the staff involved. The staff may be academic staff, or not on academic contracts, such as senior university managers or administrative/support staff.

---

**3.5.3 Primary data sources**
- Institutional accounts system or Current Research Information System (CRIS system)
- Published annual accounts
- National statutory reports, such as those available from the Higher Education Statistics Agency[93] (HESA) in the UK

**3.5.4 UK application**
Denominator derived from institutional data: institution.

Time period: financial year.

Currency: British pounds (GBP).

SOURCES OF FUNDING OF TUITION FEES FOR RESEARCH STUDENTS

ENDORSED BY: UNITED KINGDOM.

**3.6.1 Metric definition**
Research Student Funding ❄ calculates the number and proportion of research students whose tuition fees are funded by each category of funder types.

It answers the question of the sources from which research students are funded.



(a) Number of active students
(b) Proportion of active students

(a) Time period
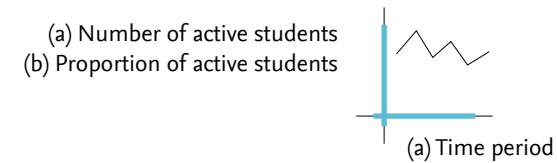
**3.6.2 Details**
- This metric is influenced by the composition of the student body, for example the mix of international and national students, and the range of disciplines.
- This metric is not concerned with the amount of funding received from each category of funder types.
- Head count, not FTE, is used as the basis for counting research students during any reporting period.

**3.6.3 Primary data sources**
- Institutional systems
- National statutory returns, such as those collected by the Higher Education Statistics Agency (HESA[94]) in the UK

---

93 www.hesa.ac.uk

94 www.hesa.ac.uk

**3.6.4 UK application**

Denominators derived from institutional data:

- HESA Joint Academic Coding System[95] (JACS) code
- Funding-type
- Full-time or part-time research students
- Home or overseas research students
- Institution

Time period: financial year.

Population: the relevant HESA standard definition should be followed to define the population for the metric calculation. The HESA standard registration population[96] forms the basis for head count for research student instances during the reporting period. The standard registration population is used as this helps determine active from non-active students; non-active students include those that are considered dormant or who are writing up. Where a student falls across multiple HESA JACS codes, the full person equivalent value should be used. It is noted that JACS will change in the near future[97], in which case the new equivalent field should be used.

**3.6.5 Future opportunities**

The amount of funding in each funding-type category would be helpful to assess the relative contribution of each funding type to the total postgraduate research student funding at an institution.

# 4. Process Metrics

This section details the methodologies for the following Snowball Metrics:

Research Metrics:
1. Income Volume
2. Market Share

Enterprise Activities / Economic Development Metrics:
3. Contract Research Volume

Post-Graduate Research Metrics:
4. Research Student to Academic Staff Ratio

---

95 www.hesa.co.uk/support/documentation/jacs

96 www.hesa.ac.uk/collection/c14051/a/feeelig/

97 JACS will be replaced by the Higher Education Costing System (HeCoS) and the Common Aggregation Hierarchy (CAH) for the 2019/20 academic year: www.hesa.ac.uk/innocation/hecos

## 4.1 Income Volume ❄

VOLUME OF RESEARCH EXPENDITURE

ENDORSED BY: UNITED KINGDOM.

### 4.1.1 Metric definition

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

It answers the question of whether the institution is undertaking more or less externally funded research than in the past.



(a) Income spent

(a) Time period

(a) Income spent per FTE

(a) Time period
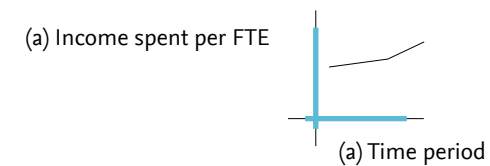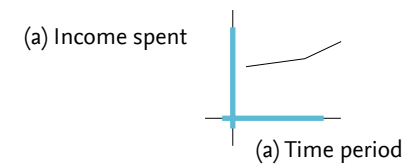
The complete Common European Research Information Format (CERIF)[98] xml code for this metric, prepared by euroCRIS, is available for download and use via the Snowball Metrics website[99].

### 4.1.2 Details

Any research that could not be classified as "Research" under HESA's accounting conventions for the finance record (Financial Reporting Standard 102)[100] will not be considered in Income Volume ❄.

98 http://www.eurocris.org/Index.php?page=featuresCERIF&t=1

99 www.snowballmetrics.com

100 https://www.hesa.ac.uk/collection/c15031/coverage#conventions

### 4.1.3 Primary data sources
- Institutional accounts system or Current Research Information System (CRIS system)
- Published annual accounts
- National statutory reports, such as those available from the Higher Education Statistics Agency[101] (HESA) in the UK

### 4.1.4 UK application
Denominators derived from institutional data:
- HESA cost centre
- Funder-type
- Institution

Time period: financial year.

Income data available from the HESA Finance Record[102] are used to generate Income Volume ✿.

Currency: British pounds (GBP).

### 4.1.5 US application
The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

The FundRef-based classification is to be used as a counterpart to the HESA funder-type mapping. See also the UK-US funder-type mapping in section 2.13.3, that enables benchmarking between UK and US institutions based on national funder-type data.

### 4.1.6 Future opportunities
A denominator reflecting themes and subject focus of income would be highly valued, especially if the same thematic denominator could be applied not only to Process, but also to Input, and Output and Outcome, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

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

---

101 www.hesa.ac.uk

102 http://www.hesa.ac.uk/index.php?option=com_pubs&task=show_pub_detail&pubid=1710&Itemid=286

## 4.2 Market Share ✿
PERCENTAGE OF SECTOR TOTAL RESEARCH INCOME PER INSTITUTION

ENDORSED BY: UNITED KINGDOM.

### 4.2.1 Metric definition
Market Share ✿ calculates the percentage of total research income across the sector related to a given institution.

It answers the questions of:
- Whether the institution is increasing or decreasing its share of available funding.
- The trend in terms of market share per funder type.



(a) Percentage of sector total research income

(a) Time period
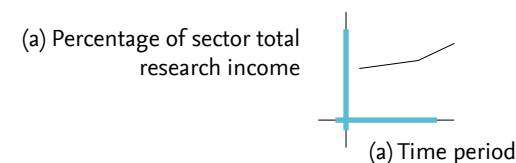
The complete Common European Research Information Format (CERIF)[103] xml code for this metric, prepared by euroCRIS, is available for download and use via the Snowball Metrics website[104].

### 4.2.2 Details
Any research that could not be classified as "Research" under HESA's accounting conventions for the finance record (Financial Reporting Standard 102)[105] will not be considered in Market Share ✿.

### 4.2.3 Primary data sources
- Institutional accounts system or Current Research Information System (CRIS system)
- Published annual accounts
- National statutory reports, such as those available from HESA in the UK

---

103 http://www.eurocris.org/Index.php?page=featuresCERIF&t=1

104 www.snowballmetrics.com

105 https://www.hesa.ac.uk/collection/c15031/coverage#conventions

**4.2.4 UK application**
Denominators derived from institutional data:
- HESA cost centre
- Funder-type
- Institution

Time period: financial year.

Income data available from the HESA Finance Record[106] are used to generate Market Share ✿. The sector total research income is the total national income as reported to HESA.

**4.2.5 US application**
The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

The FundRef-based classification is to be used as a counterpart to the HESA funder-type mapping. See also the UK-US funder-type mapping in section 2.13.3, that enables benchmarking between UK and US institutions based on national funder-type data.

**4.2.6 Future opportunities**
A denominator reflecting themes and subject focus of income would be highly valued, especially if the same thematic denominator could be applied not only to Process, but also to Input, and Output and Outcome, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

Market Share ✿ may not lend itself as easily as some other Snowball Metrics to global benchmarking, due to distinct national characteristics of competitive funding structures. It might best be viewed as a metric with multiple national flavors. Versions might be derived based on:
- Amounts awarded by funding bodies, rather than spend.
- The total available amongst participating institutions or "benchmarking clubs", rather than the national total.

## 4.3 Contract Research Volume ✿

TOTAL VALUE OF CONTRACT RESEARCH

ENDORSED BY: UNITED KINGDOM.

**4.3.1 Metric definition**
Contract Research Volume ✿ calculates the value of income from contract research.

It answers the question of how much an institution is doing to address the needs of industry.



(a) Value of income
(a) Time period

(a) Value of income per FTE
(a) Time period
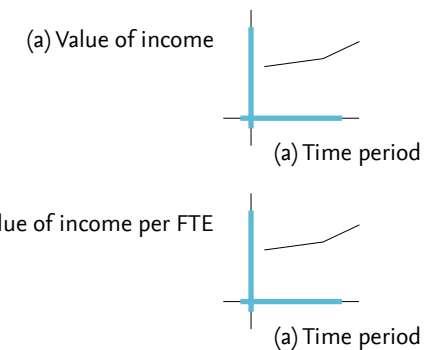
**4.3.2 Details**
Contract research income is that received from an industrial or private external body, which is neither a university nor an academic organization, from commissioning a particular piece of research with specific terms. Contract research is targeted at solving a particular problem, whereas other research questions may be more loosely defined and focused on discovery. The information and intellectual property arising from contract research will contractually be at least partly owned by the third party that is paying for the work.

---

106 http://www.hesa.ac.uk/index.php?option=com_pubs&task=show_pub_detail&pubid=1710&Itemid=286

Contract research volume excludes:
- Research income from external funding bodies, which is covered in Awards Volume ❄.
- Income associated with research projects that are publicly sponsored, and that are performed in collaboration with at least one other non-academic organization, which is covered in Academic-Industry Leverage ❄.

**4.3.3 Primary data sources**
- Institutional accounts system or Current Research Information System (CRIS system)
- Published annual accounts
- National statutory reports, such as those available from the Higher Education Statistics Agency[107] (HESA) in the UK

**4.3.4 UK application**
Denominator derived from institutional data: institution.

Time period: financial year.

Currency: British pounds (GBP).

## 4.4 Research Student to Academic Staff Ratio ❄
RATIO OF RESEARCH STUDENTS TO ACADEMIC STAFF INVOLVED IN RESEARCH

ENDORSED BY: UNITED KINGDOM.

**4.4.1 Metric definition**
Research Student to Academic Staff Ratio ❄ calculates the average number of research students per researcher (see section 2.10 for definition of researcher).

It answers the questions of:
- The number of research students that academics attract and supervise.
- The attractiveness of the institution for research students and relevant funding.
- Which disciplines attract the highest proportion of research students.



(a) Ratio of research students : researchers

(a) Time period
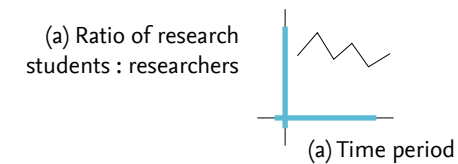
**4.4.2 Details**
- This metric is concerned with researchers who are eligible to supervise research students. For the sake of clarity, this metric is not concerned with undergraduate students.
- The "target" for this metric is not similar across institutions, and therefore not necessarily 1. The situation of the particular institution will determine whether a high or low value of this ratio is desired, and should be known when using this metric for benchmarking to enable a proper interpretation. For instance, the ratio for a research-intensive institution may be higher than for a teaching-focused institution. The spread of disciplines will also affect this metric at institutional level; it will typically be a fraction in arts and humanities.

---

**4.4.3 Primary data sources**
- Institutional systems
- National statutory returns, such as those collected by the Higher Education Statistics Agency (HESA[108]) in the UK

**4.4.4 UK application**
Denominators derived from institutional data:
- HESA cost centre
- Institution

Time period: financial year.

Population: the relevant HESA standard definition should be followed to define the population for the metric calculation.

**4.4.5 Future opportunities**
It would be beneficial to consider using the number of supervisors, rather than the number of those eligible to supervise, in the recipe to better reflect the overall attractiveness of the research environment.

It would be beneficial to adjust for franchised-out or collaborative provisions. For example, when a student is being supervised 50% by two universities, as part of a collaboration, the adjustment could be used to calculate the ratio that the student is taught or supervised at each institution.

---

# 5. Output and Outcome Metrics

This section details the methodologies for the following Snowball Metrics:

**Research Metrics:**
- Publications and Citations
  1. Scholarly Output
  2. Citation Count
  3. Citations per Output
  4. *h*-index
  5. Field-Weighted Citation Impact
  6. Outputs in Top Percentiles
  7. Publications in Top Journal Percentiles

- Collaboration
  8. Collaboration
  9. Collaboration Publication Share
  10. Collaboration Impact
  11. Collaboration Field-Weighted Citation Impact
  12. Academic-Corporate Collaboration
  13. Academic-Corporate Collaboration Impact

- Societal impact
  14. Altmetrics
  15. Public Engagement
  16. Academic Recognition

**Enterprise Activities / Economic Development Metrics:**
  17. Intellectual Property Volume
  18. Intellectual Property Income
  19. Sustainable Spin-Offs
  20. Spin-Off-Related Finances

**Post-Graduate Research Metrics:**
  21. Time to Award of Doctoral Degree
  22. Destination of Research Student Leavers

## 5.1 Scholarly Output ✿

PRODUCTIVITY BASED ON ANY TYPE OF SCHOLARLY OUTPUT

ENDORSED BY: UNITED KINGDOM.

### 5.1.1 Metric definition
Scholarly Output ✿ counts the number of institutional outputs of any type.

It answers the question of how productive an institution is in generating scholarly outputs, such as publications.



(a) Number of outputs / (a) Time period

(a) Number of outputs per FTE / (a) Time period
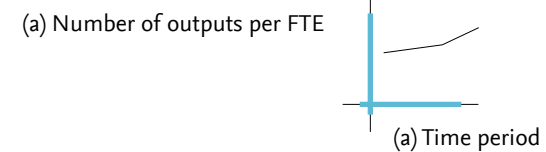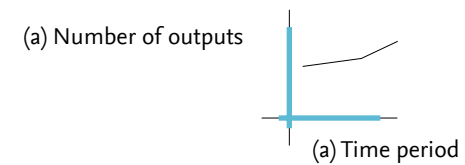
The complete Common European Research Information Format (CERIF)[109] xml code for this metric, prepared by euroCRIS, is available for download and use via the Snowball Metrics website[110].

---

109 http://www.eurocris.org/Index.php?page=featuresCERIF&t=1
110 www.snowballmetrics.com

**5.1.2 Details**
Outputs of the following types are included in Scholarly Output ❄:
- Journal publications
- Book series
- Stand-alone books, defined as:
  - Edited volumes or anthologies
  - Monographs or scholarly editions, including scholarly biographies
  - Major Reference Works where the items include cited references
  - Text books aimed at a graduate or advanced research audience
- Artefacts
- Compositions
- Designs
- Devices and Products
- Digital or visual media
- Exhibitions
- Internet publications
- Performances
- Reports, whether confidential, technical or commissioned
- Software

Outputs of the following types are excluded from Scholarly Output ❄:
- Patents (they are counted in Intellectual Property Volume ❄)
- Theses (these may be addressed separately in future in the post-graduate research metrics)

Scholarly Output ❄ defines the total count of items, to represent productivity. It may be useful to be able to count and/or exclude sub-groups from the total count; these sub-sets will not be defined by Snowball Metrics, but left to the discretion of the implementer.

**5.1.3 Primary data sources**
- Institutional output repository or Current Research Information System (CRIS system)
- Scopus
- Web of Science / Book Citation Index
- Google Scholar
- WorldCat[111]

**5.1.4 UK application**
Denominators derived from institutional data:
- HESA cost centre. Outputs are associated with a HESA cost centre via the researcher(s) who produced them.
- Institution

Time period: calendar year.

**5.1.5 US application**
The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

**5.1.6 Future opportunities**
Commercial abstracting and indexing databases continue to 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 reflecting the themes and subject focus of outputs would be highly valued, especially if the same thematic denominator could be applied not only to Output and Outcome, but also to Input and Process, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

It is desirable in the future to add research students as an additional denominator to this recipe, but the Post-Graduate Research expert group concluded that this is not possible from institutional data sources at the moment.

## 5.2 Citation Count ❄

TOTAL CITATIONS RECEIVED BY INSTITUTION'S SCHOLARLY OUTPUT

ENDORSED BY: UNITED KINGDOM.

### 5.2.1 Metric definition
Citation Count ❄ sums the citations received to date by institutional outputs.

It answers the question of how many total citations an institution's output has received, as a proxy for the academic impact of these outputs.

The complete Common European Research Information Format (CERIF)[112] xml code for this metric, prepared by euroCRIS, is available for download and use via the Snowball Metrics website[113].

(a) Number of citations

(a) Time period

(a) Number of citations per FTE

(a) Time period
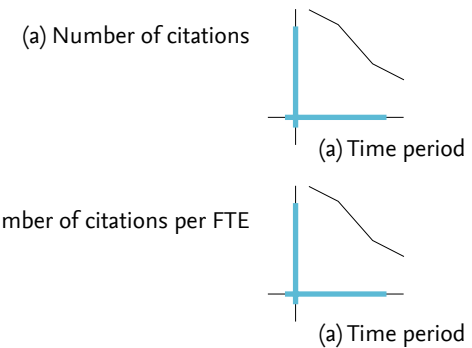
### 5.2.2 Details
It is likely that citation data will not be available for all elements that constitute an institution's Scholarly Output ❄. For example, if a commercial abstracting and indexing database is used as the data source for Citation Count ❄, their coverage will be less than 100% of the institution's total productivity. A partial reflection of an institution's activity is still valuable in providing an evidence-based support for decision making through benchmarking, since this limitation is likely to affect all comparators equally.

---

112 http://www.eurocris.org/Index.php?page=featuresCERIF&t=1
113 www.snowballmetrics.com

**5.2.3 Primary data sources**
- Scopus
- Web of Science
- Google Scholar

**5.2.4 UK application**
Denominators derived from institutional data:
- HESA cost centre. Outputs are associated with a HESA cost centre via the researcher(s) who produced them
- Institution

Time period: calendar year. Note that the time period does not refer to the year in which citations were received, but to the year in which outputs were produced.

**5.2.5 US application**
The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

**5.2.6 Future opportunities**
Commercial abstracting and indexing databases continue to 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 reflecting the themes and subject focus of outputs would be highly valued, especially if the same thematic denominator could be applied not only to Output and Outcome, but also to Input and Process, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

## 5.3 Citations per Output ❄

AVERAGE CITATIONS RECEIVED BY AN ITEM OF SCHOLARLY OUTPUT

ENDORSED BY: UNITED KINGDOM.

**5.3.1 Metric definition**
Citations per Output ❄ calculates the average citations received to date by each output that is part of a particular set.

It answers the question of how many citations an institution's outputs have each received on average, as a proxy for the academic impact of these outputs.

(a) Number of citations per output

(a) Time period
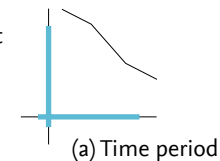
**5.3.2 Details**
It is likely that citation data will not be available for all elements that constitute an institution's Scholarly Output ❄. For example, if a commercial abstracting and indexing database is used as the data source for Citations per Output ❄, their coverage will be less than 100% of 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 through benchmarking, since this limitation is likely to affect all comparators equally.

**5.3.3 Primary data sources**
- Scopus
- Web of Science
- Google Scholar

### 5.3.4 UK application

Denominators derived from institutional data:

- HESA cost centre. Outputs are associated with a HESA cost centre via the researcher(s) who produced them
- Institution

Time period: calendar year. Note that the time period does not refer to the year in which citations were received, but to the year in which outputs were produced.

### 5.3.5 US application

The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

### 5.3.6 Future opportunities

Commercial abstracting and indexing databases continue to 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 reflecting the themes and subject focus of outputs would be highly valued, especially if the same thematic denominator could be applied not only to Output and Outcome, but also to Input and Process, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

## 5.4 *h*-index ✿

*H*-INDEX PER DISCIPLINE

ENDORSED BY: UNITED KINGDOM.

### 5.4.1 Metric definition

The *h*-index ✿ is calculated, as defined by Professor Jorge Hirsch[114], 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 $Np$ papers have at least $h$ citations each and the other $(Np - h)$ 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.

It answers the question of what is the depth of quality of the outputs of a group of academics using citation counts as a proxy for the quality of these outputs.

The *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.



(a) *h*-index

(a) Discipline
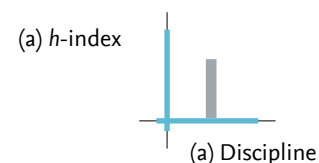
The complete Common European Research Information Format (CERIF)[115] xml code for this metric, prepared by euroCRIS, is available for download and use via the Snowball Metrics website[116].

114 Hirsch, J. E. (2005). "An index to quantify an individual's scientific research output" Proc. Natl Acad. Sci. 201 (46): 16569–16572.

115 http://www.eurocris.org/Index.php?page=featuresCERIF&t=1

116 www.snowballmetrics.com

### 5.4.2 Details

The *h*-index ✿ is not calculated at institutional level. The *h*-index ✿ was originally conceived of as a useful reflection of a researcher's accumulated career, and is represented by a single number which stays the same or increases with time, but which cannot go down. One of the aims of Snowball Metrics is to understand current excellence in recent university performance, and as such metrics values must be able to fall as well as increase: *h*-index ✿ does not encompass this possibility. The project partners conceive of using *h*-index ✿ as it was intended, for researcher-related, disciplinary denominators only.

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 the *h*-index ✿, 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 through benchmarking, since this limitation is likely to affect all comparators equally.

### 5.4.3 Primary data sources
- Scopus
- Web of Science
- Google Scholar

### 5.4.4 UK application
Denominator derived from institutional data: HESA cost centre, via assignment of a researcher associated with an output to a HESA cost centre.

### 5.4.5 US application
The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

### 5.4.6 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 reflecting the themes and subject focus of outputs would be highly valued, especially if the same thematic denominator could be applied not only to Output and Outcome, but also to Input and Process, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

## 5.5 Field-Weighted Citation Impact

ACTUAL CITATION COUNT RELATIVE TO THE EXPECTED WORLD CITATION COUNT

ENDORSED BY: UNITED KINGDOM.

### 5.5.1 Metric definition

Field-Weighted Citation Impact ❄ is the ratio of the citations actually received by the denominator's output, and the average number of citations received by all other similar publications. A Field-Weighted Citation Impact ❄ of:

- Exactly 1.00 means that the output performs just as expected for the global average.
- More than 1.00 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; for example, 0.91 means 9% less cited than expected.

It answers the question of whether the institution's outputs are cited above or below 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, or when comparing between 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.
- If these differences are not accounted for:
  - The effects of outputs in medicine and biochemistry will dominate the effects of those in mathematics and education, in a denominator comprising multiple disciplines, such as an institution.
  - Different levels of performance will be disguised by these characteristic behavioral traits when comparing between disciplines.
- This means that, if using non-weighted metrics, an institution that is focused on medicine will generally appear to perform better than an institution that specializes in social sciences.
- The methodology of Field-Weighted Citation Impact ❄ accounts for these disciplinary differences

(a) Field-Weighted Citation Impact

(a) Time period
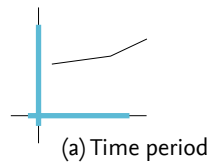
The complete Common European Research Information Format (CERIF)[117] xml code for this metric, prepared by euroCRIS, is available for download and use via the Snowball Metrics website[118].

**5.5.2 Details**
The expected average 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. This is an exception to the general approach of counting total citations received since publication. For example:
- For an item published in October 2007, citations will be counted until the end of December 2010.
- For an item published in June 2015, citations will be counted until the end of December 2018. If calculating this metric before December 2018, this complete 3-year window cannot be used; in this situation, the citations will be counted up to the current date.

If an output is part of more than one subject field:
- Its publication and citation counts are distributed equally across all subject fields of which it is part, so that a single output does not exert too much weight. If an output is part of 2 subject fields, it is counted as 0.5 outputs per subject field, and the citations it has received are shared equally between them.
- The expected citations per output for each field are determined, and the harmonic average[119] 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 in the time period 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 through benchmarking, since this limitation is likely to affect all comparators equally.

Field-Weighted Citation Impact is steadier over time for a larger set of publications. It may be useful to implement this recipe for multiples of publication years, if fluctuation is a concern. It is recommended considering Scholarly Output ✱ alongside this metric to communicate a better understanding of the cause of any fluctuations.

**5.5.3 Primary data sources**
- Scopus
- Web of Science
- Google Scholar

**5.5.4 UK application**
Denominators derived from institutional data:
- HESA cost centre. Outputs are associated with a HESA cost centre via the researcher(s) who produced them.
- Institution

Time period: calendar year. Note that the time period does not refer to the year in which citations were received, but to the year in which outputs were produced.

---

117 http://www.eurocris.org/Index.php?page=featuresCERIF&t=1

118 www.snowballmetrics.com

119 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.

### 5.5.5 US application

The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

### 5.5.6 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 reflecting the themes and subject focus of outputs would be highly valued, especially if the same thematic denominator could be applied not only to Output and Outcome, but also to Input and Process, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

## 5.6 Outputs in Top Percentiles ✿

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

ENDORSED BY: UNITED KINGDOM, UNITED STATES.

### 5.6.1 Metric definition

The citation thresholds that represent the top 1%, 5%, 10% and 25% of 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.

It answers the question of what number or fraction of an institution's outputs have reached a particular citation threshold in the data universe.
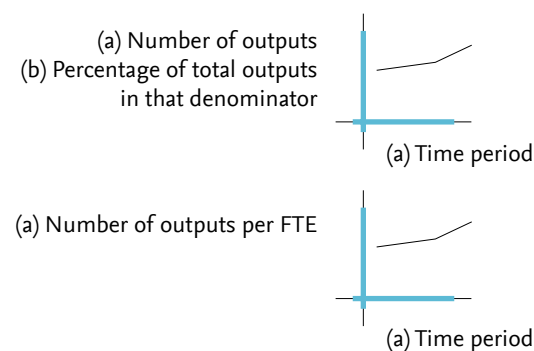


The complete Common European Research Information Format (CERIF)[120] xml code for this metric, prepared by euroCRIS, is available for download and use via the Snowball Metrics website[121].

---

120 http://www.eurocris.org/Index.php?page=featuresCERIF&t=1
121 www.snowballmetrics.com

### 5.6.2 Details

The metric Outputs in Top Percentiles ❄ depends on being able to divide outputs in the data universe into 100 percentiles. Early in the current calendar year, it is unlikely that enough citations will have been received by outputs to enable this, especially at more granular denominators such as disciplines. This metric will only be calculable some months into the current year, and we suggest from 1 July.

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 through benchmarking, since this limitation is likely to affect all comparators equally.

### 5.6.3 Primary data sources
- Institutional output repository and Current Research Information System (CRIS system)
- Scopus
- Web of Science
- Google Scholar

### 5.6.4 UK application
Denominators derived from institutional data:
- HESA cost centre, via assignment of a researcher associated with an output to a HESA cost centre.
- Institution

Time period:
- Calendar year
- Rolling 3-year blocks

Note that the time period does not refer to the year in which citations were received, but to the year in which outputs were produced.

### 5.6.5 US application
The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

### 5.6.6 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 reflecting the themes and subject focus of outputs would be highly valued, especially if the same thematic denominator could be applied not only to Output and Outcome, but also to Input and Process, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

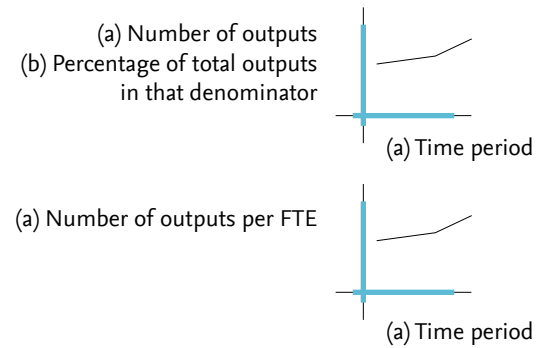## 5.7 Publications in Top Journal Percentiles ❅

OUTPUTS THAT HAVE BEEN PUBLISHED IN SERIALS WITH A PARTICULAR AVERAGE CITATION
THRESHOLD IN THE DATA UNIVERSE

ENDORSED BY: UNITED KINGDOM.

### 5.7.1 Metric definition
The average citation thresholds that represent the top 1%, 5%, 10% and 25% of
serials 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.

It answers the question of what number or fraction of an institution's outputs
have been published in serials with a particular average citation threshold in the
data universe.

(a) Number of outputs
(b) Percentage of total outputs
in that denominator

(a) Time period

(a) Number of outputs per FTE

(a) Time period

### 5.7.2 Details

The metric Publications in Top Journal Percentiles ❄ depends on being able to divide the serials indexed by the data universe into 100 percentiles. Any journal metric that enables this can be employed in this metric, and this metric could have multiple versions depending on which journal metric is used to create percentiles from the data universe.

The percentile boundaries are calculated independently for each publication year, not overall for the entire data universe, and an output is compared to the percentile boundaries for its publication year. For example:

- An output published in 2008 is counted in the top 1% if it is part of serials that are in the top 1% of the data universe according to 2008 journal metrics. It is irrelevant if these serials are no longer part of the top 1% in 2009.
- For recently published outputs, the relevant publication year's journal ranking metric may not yet be available; 2017 journal metrics will be released in 2018, for instance. In this situation, the journal metric for the previous calendar year should be used until the journal metrics for the actual publication year become available.
- Some journal metric values may not be available for older outputs; CiteScore is available from 2011, for instance, and Source-Normalized Impact per Paper (SNIP) and SCImago Journal Rank (SJR) are available from 1999 onwards. In this instance, for outputs published in serials before the first available journal metric date, the oldest journal metric value is used.

This metric is not itself field-normalized, although if the journal metric on which it is based takes different behavior between disciplines into account, then it will behave as a field-normalized metric. Of those journal metrics mentioned in the "Primary data sources" section, CiteScore and the Impact Factor® are not field-normalized; the tendency for life sciences journals to have higher average citations per output than, say, social sciences journals, might affect the choice of denominators for which a CiteScore- or Impact Factor®-based Publications in Top Journal Percentiles ❄ is used, whereas this might not be a consideration when employing the other journal metrics to generate this metric.

It is likely that journal ranking metrics will not be available for all elements that constitute an institution's Scholarly Output ❄. For example, stand-alone books cannot, by definition, have an Eigenfactor, or equivalent, because these can only be calculated for serial publications. A partial reflection of an institution's activity is still valuable in providing an evidence-based support for decision making through benchmarking, since this limitation is likely to affect all comparators equally.

### 5.7.3 Primary data sources

- Institutional output repository and Current Research Information System (CRIS system)
- Scopus
- Web of Science
- Google Scholar
- Journal ranking metrics such as Impact Factor® [122], and the following journal metrics that are available free-of-charge: CiteScore metrics, Source-Normalized Impact per Paper (SNIP) and SCImago Journal Rank (SJR)[123]; and Eigenfactor and Article Influence[124].

### 5.7.4 UK application

Denominators derived from institutional data:

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

Time period: calendar year. Note that the time period does not refer to the year in which citations were received, but to the year in which outputs were produced.

### 5.7.5 US application

The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

---

122 http://thomsonreuters.com/journal-citation-reports/
123 www.journalmetrics.scopus.com
124 www.eigenfactor.org

### 5.7.6 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 reflecting the themes and subject focus of outputs would be highly valued, especially if the same thematic denominator could be applied not only to Output and Outcome, but also to Input and Process, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

## 5.8 Collaboration ✼

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

ENDORSED BY: UNITED KINGDOM, UNITED STATES.

### 5.8.1 Metric definition

Collaboration ✼ calculates the number and 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, to avoid double counting.
- Countries are defined as in the ISO classification.[125]

It answers the question of what proportion of the institution's publications are produced in collaboration with national versus international partners.

(a) Number of internationally collaborative outputs
(b) Internationally collaborative outputs as percentage of total outputs in that denominator
(c) Number of nationally collaborative outputs
(d) Nationally collaborative outputs as percentage of total outputs in that denominator

(a) Time period

(a) Number of internationally collaborative outputs per FTE
(b) Number of nationally collaborative outputs per FTE

(a) Time period

The complete Common European Research Information Format (CERIF)[126] xml code for this metric, prepared by euroCRIS, is available for download and use via the Snowball Metrics website[127].

---

125 http://www.iso.org/iso/ics6-en.pdf
126 http://www.eurocris.org/Index.php?page=featuresCERIF&t=1
127 www.snowballmetrics.com

### 5.8.2 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 centers. 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:

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

The country information specified 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 through benchmarking, since this limitation is likely to affect all comparators equally.

### 5.8.3 Primary data sources

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

- Institutional output repository and Current Research Information System (CRIS system)
- Scopus
- Web of Science
- Google Scholar

### 5.8.4 UK application

Denominators derived from institutional data:

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

Time period: calendar year. Note that the time period does not refer to the year in which citations were received, but to the year in which outputs were produced.

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

### 5.8.5 US application

The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

### 5.8.6 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 reflecting the themes and subject focus of outputs would be highly valued, especially if the same thematic denominator could be applied not only to Output and Outcome, but also to Input and Process, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

## 5.9 Collaboration Publication Share ❄

PROPORTION OF INSTITUTIONAL OUTPUT ASSOCIATED WITH A PARTICULAR
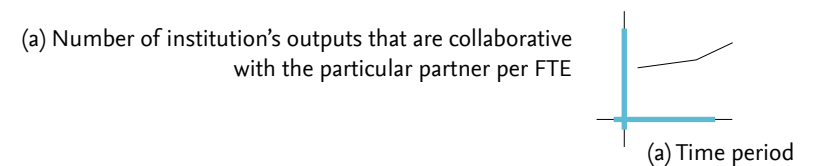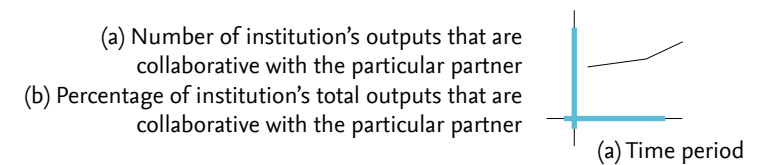COLLABORATION PARTNER

ENDORSED BY: UNITED KINGDOM.

### 5.9.1 Metric definition

Collaboration Publication Share ❄ calculates the proportion of institutional
output that is associated with a particular collaboration partner.

Any collaboration partner of interest to the institution can be considered,
for example:
- A country.
- A set of countries related by social, political or economic considerations such
  as the BRIC countries: Brazil, Russia, India and China, which are all deemed
  to be at a similar stage of newly advanced economic development.
- Another institution.
- A group of institutions, such as the Russell Group in the UK, the Ivy League
  in the US, or the set of peers that an institution has chosen to benchmark
  itself against

It answers the question of the extent to which an institution collaborates with a
particular collaboration partner.

(a) Number of institution's outputs that are
collaborative with the particular partner
(b) Percentage of institution's total outputs that are
collaborative with the particular partner

(a) Time period

(a) Number of institution's outputs that are collaborative
with the particular partner per FTE

(a) Time period

**5.9.2 Details**

Single author papers are excluded in this recipe.

The country information specified 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.

An output could be counted as part of multiple Collaboration Publication Shares ❄ if it has more than 2 affiliations; the affiliation of the institution in question, and at least 2 others for which it forms part of the collaboration output.

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 through benchmarking, since this limitation is likely to affect all comparators equally.

**5.9.3 Primary data sources**
- Scopus
- Web of Science
- Google Scholar

**5.9.4 UK application**
Denominators derived from institutional data:
- HESA cost centre. Outputs are associated with a HESA cost centre via the researcher(s) who produced them.
- Institution

Time period: calendar year. Note that the time period does not refer to the year in which citations were received, but to the year in which outputs were produced.

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

**5.9.5 US application**
The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

**5.9.6 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 reflecting the themes and subject focus of outputs would be highly valued, especially if the same thematic denominator could be applied not only to Output and Outcome, but also to Input and Process, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

## 5.10 Collaboration Impact ❄

CITATION IMPACT OF NATIONALLY AND INTERNATIONALLY CO-AUTHORED
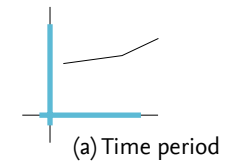SCHOLARLY OUTPUTS

ENDORSED BY: UNITED KINGDOM.

### 5.10.1 Metric definition

Collaboration Impact ❄ calculates the average citations received by the sets of
output 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, to avoid double counting.
- Countries are defined as in the ISO classification.[128]

It answers the question of the citation impact of an institution's nationally versus
internationally collaborative publications, as a proxy for the academic impact of
these outputs.

(a) Citations per internationally collaborative output
    (b) Citations per nationally collaborative output

(a) Time period

---

128 http://www.iso.org/iso/ics6-en.pdf

**5.10.2 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 centers. 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:

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

The country information specified 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.

The assignment of international or national applies only to the institutional outputs. The count of citations is independent of the international or national collaboration status of the citing output; if an institution's internationally collaborative output has only been cited by nationally collaborative publications, then these citations are still counted.

It is likely that affiliation and / or 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 the collaboration information, and 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 through benchmarking, since this limitation is likely to affect all comparators equally.

**5.10.3 Primary data sources**

- Scopus
- Web of Science
- Google Scholar

**5.10.4 UK application**

Denominators derived from institutional data:

- HESA cost centre. Outputs are associated with a HESA cost centre via the researcher(s) who produced them.
- Institution

Time period: calendar year. Note that the time period does not refer to the year in which citations were received, but to the year in which outputs were produced.

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

**5.10.5 US application**

The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

**5.10.6 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 reflecting the themes and subject focus of outputs would be highly valued, especially if the same thematic denominator could be applied not only to Output and Outcome, but also to Input and Process, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

## 5.11 Collaboration Field-Weighted Citation Impact ❄

FIELD-WEIGHTED CITATION IMPACT ❄ OF INSTITUTIONAL OUTPUT ASSOCIATED
WITH A PARTICULAR COLLABORATION PARTNER

ENDORSED BY: UNITED KINGDOM.

### 5.11.1 Metric definition

Collaboration Field-Weighted Citation Impact ❄ calculates the Field-Weighted
Citation Impact ❄ of institutional output that is associated with a particular
collaboration partner.

Any collaboration partner of interest to the institution can be considered,
for example:
* A country.
* A set of countries related by social, political or economic considerations such
  as the BRIC countries: Brazil, Russia, India and China, which are all deemed
  to be at a similar stage of newly advanced economic development.
* Another institution.
* A group of institutions, such as the Russell Group in the UK, the Ivy League
  in the US, or the set of peers that an institution has chosen to benchmark
  itself against

(a) Field-Weighted Citation Impact ❄
    of an institution's outputs that are
collaborative with the particular partner

(a) Time period
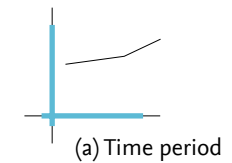
It answers the question of the extent to which an institution's output is more or
less cited through collaboration with a particular collaboration partner

### 5.11.2 Details

The country information specified 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.

The assignment of international or national applies only to the institutional outputs. The count of citations is independent of the international or national collaboration status of the citing output; if an institution's internationally collaborative output has only been cited by nationally collaborative publications, then these citations are still counted.

It is likely that affiliation and / or 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 the collaboration information, and 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 through benchmarking, since this limitation is likely to affect all comparators equally.

Collaboration Field-Weighted Citation Impact ✿ (and Field-Weighted Citation Impact ✿) is steadier over time for a larger set of publications. It may be useful to implement this recipe for multiples of publication years, if fluctuation is a concern. It is recommended considering Collaboration Publication Share ✿ alongside this metric to communicate a better understanding of the cause of any fluctuations.

### 5.11.3 Primary data sources
- Scopus
- Web of Science
- Google Scholar

### 5.11.4 UK application

Denominators derived from institutional data:
- HESA cost centre. Outputs are associated with a HESA cost centre via the researcher(s) who produced them.
- Institution

Time period: calendar year. Note that the time period does not refer to the year in which citations were received, but to the year in which outputs were produced.

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

### 5.11.5 US application

The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

### 5.11.6 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 reflecting the themes and subject focus of outputs would be highly valued, especially if the same thematic denominator could be applied not only to Output and Outcome, but also to Input and Process, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.
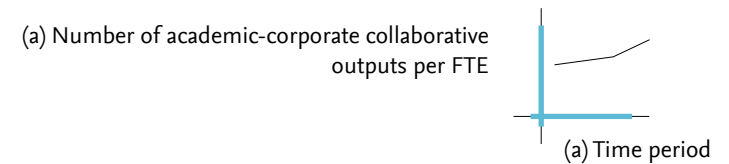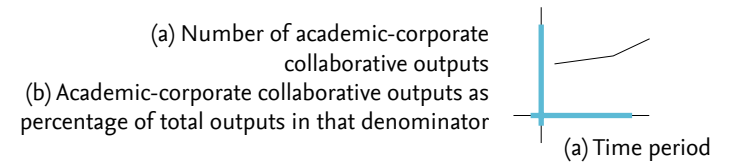
## 5.12 Academic-Corporate Collaboration ❄

VOLUME AND PROPORTION OF SCHOLARLY OUTPUTS CO-AUTHORED BY RESEARCHERS FROM BOTH ACADEMIC AND CORPORATE AFFILIATIONS

ENDORSED BY: UNITED KINGDOM.

### 5.12.1 Metric definition
Academic-Corporate Collaboration ❄ calculates the number and percentage of outputs that have been co-authored by researchers from both academic and corporate, or industrial, affiliations.

It answers the question of what proportion of the institution's publications are produced in collaboration with corporate partners.



(a) Number of academic-corporate collaborative outputs
(b) Academic-corporate collaborative outputs as percentage of total outputs in that denominator

(a) Time period

(a) Number of academic-corporate collaborative outputs per FTE

(a) Time period

### 5.12.2 Details
It is likely that affiliation and affiliation-type 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, its 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 through benchmarking, since this limitation is likely to affect all comparators equally.

### 5.12.3 Primary data sources

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

- Institutional output repository and Current Research Information System (CRIS system)
- Scopus
- Web of Science
- Google Scholar

### 5.12.4 UK application

Denominators derived from institutional data:

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

Time period: calendar year. Note that the time period does not refer to the year in which citations were received, but to the year in which outputs were produced.

### 5.12.5 US application

The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2,9,3, that enables benchmarking between UK and US institutions based on national discipline data.

### 5.12.6 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 reflecting the themes and subject focus of outputs would be highly valued, especially if the same thematic denominator could be applied not only to Output and Outcome, but also to Input and Process, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.
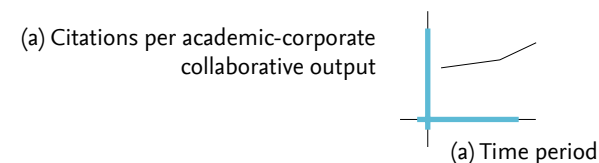
## 5.13 Academic-Corporate Collaboration Impact ❄

CITATION IMPACT OF SCHOLARLY OUTPUTS CO-AUTHORED BY RESEARCHERS FROM BOTH ACADEMIC AND CORPORATE AFFILIATIONS

ENDORSED BY: UNITED KINGDOM.

### 5.13.1 Metric definition

Academic-Corporate Collaboration Impact ❄ calculates the average citations received by the outputs that have been co-authored by researchers from both academic and corporate or industrial, affiliations.

It answers the question of the citation impact of an institution's publications which have been co-authored with corporate partners, as a proxy for the academic impact of these outputs.



(a) Citations per academic-corporate collaborative output

(a) Time period

### 5.13.2 Details

The assignment or not of an academic-corporate collaborative output applies only to the institutional outputs. The count of citations is independent of the collaboration status of the citing output; if an institution's academic-corporate collaborative output has only been cited by publications authored solely by researchers with academic affiliations, then these citations are still counted.

It is likely that affiliation, affiliation-type and / or 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 the collaboration information, and 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 through benchmarking, since this limitation is likely to affect all comparators equally.

### 5.13.3 Primary data sources
- Scopus
- Web of Science
- Google Scholar

### 5.13.4 UK application
Denominators derived from institutional data:
- HESA cost centre. Outputs are associated with a HESA cost centre via the researcher(s) who produced them.
- Institution

Time period: calendar year. Note that the time period does not refer to the year in which citations were received, but to the year in which outputs were produced.

### 5.13.5 US application
The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

### 5.13.6 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 reflecting the themes and subject focus of outputs would be highly valued, especially if the same thematic denominator could be applied not only to Output and Outcome, but also to Input and Process, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

## 5.14 Altmetrics ✳
ONLINE ACTIVITY STIMULATED BY SCHOLARLY OUTPUT

ENDORSED BY: UNITED KINGDOM.

### 5.14.1 Metric definition
Altmetrics ✳ counts the number of online events that have been stimulated by an institution's output. This metric divides the broad range of online events into 4 categories:

- **Scholarly Activity** refers to the number of times that an institution's output has been posted in online tools that are typically used by academic scholars, such as Mendeley[129], CiteULike[130], Google Scholar Library[131], QUOSA[132], Papers[133], ScienceScape[134], MyScienceWork[135], colwiz[136], zotero[137], Academia.edu[138], ResearchGate[139], VIVO[140], and Scopus' "Add to My List" application.[141]
  - It answers the question of how often an institution's output is being downloaded, bookmarked or captured in online tools typically used by academic scholars, when there is an intention to come back to use the output at a later date.

- **Scholarly Commentary** refers to the number of times that an institution's output has been commented on in online tools that are typically used by academic scholars, such as science blogs, video posts such as those on YouTube[142] and vimeo[143], peer reviews such as Publons[144], post-publication comments such as PubMed Commons[145], Faculty of 1000[146] reviews, Stack Exchange[147], and Wikipedia[148] posts and citations.
  - It answers the question of how often an institution's output is being mentioned in online tools typically used by academic scholars to discuss, engage or review research outputs.

---

129 www.mendeley.com
130 www.citeulike.org
131 http://scholar.google.com
132 www.elsevier.com/online-tools/quosa
133 www.papersapp.com/mac
134 https://sciencescape.org
135 www.mysciencework.com

136 www.colwiz.com
137 www.zotero.org
138 www.academia.edu
139 www.researchgate.net
140 www.vivoweb.org
141 www.scopus.com
142 www.youtube.com

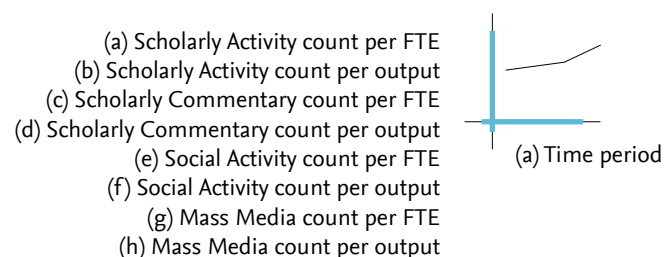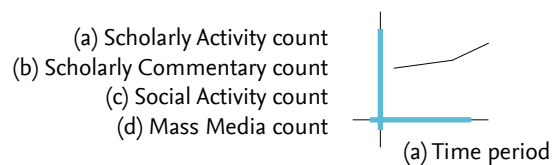143 https://vimeo.com
144 https://publons.com
145 http://www.ncbi.nlm.nih.gov/pubmedcommons
146 http://f1000.com
147 http://stackexchange.com
148 www.wikipedia.org

- **Social Activity** refers to the number of times that an institution's output has stimulated social media posts, such as those on Facebook[149],Twitter[150], Reddit[151], Google+[152], Pinterest[153], LinkedIn[154], and delicio.us.[155]
  - It answers the questions of:
    - How often an institution's output has stimulated attention or buzz in social networking services.
    - How well the institution's researchers are promoting their work through social activity.

- **Mass Media** refers to the number of times that an institution's output has been referred to by press clippings and news websites, such as The Guardian[156] newspaper.
  - It answers the question of how often the institution's output is mentioned in mass media outlets.



(a) Scholarly Activity count
(b) Scholarly Commentary count
(c) Social Activity count
(d) Mass Media count

(a) Time period



(a) Scholarly Activity count per FTE
(b) Scholarly Activity count per output
(c) Scholarly Commentary count per FTE
(d) Scholarly Commentary count per output
(e) Social Activity count per FTE
(f) Social Activity count per output
(g) Mass Media count per FTE
(h) Mass Media count per output

(a) Time period

The field of altmetrics is still new and dynamic, and much research is being conducted. In addition, the online sources that have been indexed and that are thus available to contribute to the counts of online activity are evolving. For this reason, Altmetrics ❄ defines standard "buckets" within which the various indexed online sources can be distributed. The online tools listed above, within this section, should not be taken as an exhaustive list, but rather as examples to illustrate the type of activity that should be counted in each category.

### 5.14.2 Details

Information contained within the online activity is used to attribute it to an institution's output. Depending on the data source used, this may be done by one or more of the following: resolving a DOI®[157] (Digital Object Identifier), resolving a shortened DOI such as by using bitly[158], or parsing data, perhaps semi-manually, by detecting pattern strings.

It is likely that altmetric data will not be available for all elements that constitute an institution's Scholarly Output ❄. For example, it may not be possible to resolve online activity to exhibitions or performances, and mass media often does not mention the output on which they are basing a piece in a way that can be automatically recognized, so that the database 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 through benchmarking, since this limitation is likely to affect all comparators equally.

### 5.14.3 Primary data sources
- Scopus
- Web of Science
- Google Scholar
- Indexes of online activity that can be resolved to particular individual outputs, such as Altmetric[159], Public Library of Science[160], ImpactStory[161], and Plum Analytics[162] which is owned by Elsevier[163].

### 5.14.4 UK application
Denominators derived from institutional data:
- HESA cost centre. Outputs are associated with a HESA cost centre via the researcher(s) who produced them.
- Institution

Time period: calendar year. Note that the time period does not refer to the year in which online activity was recorded, but to the year in which outputs were produced.

149 www.facebook.com
150 https://twitter.com
151 www.reddit.com
152 https://plus.google.com
153 www.pinterest.com
154 www.linkedin.com
155 https://delicious.com
156 www.theguardian.com/uk

157 www.doi.org
158 https://bitly.com
159 www.altmetric.com
160 www.plos.org
161 https://impactstory.org
162 www.plumanalytics.com
163 www.ebsco.com

### 5.14.5 US application
The HERD mapping is to be used, as the counterpart to the UK HESA cost centre mapping. See also the HESA-HERD mapping in the section 2.9.3, that enables benchmarking between UK and US institutions based on national discipline data.

### 5.14.6 Future opportunities
Additional categories of Altmetrics ❄ may be added in future editions of this recipe book.

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 reflecting the themes and subject focus of outputs would be highly valued, especially if the same thematic denominator could be applied not only to Output and Outcome, but also to Input and Process, metrics. Most likely, an automated way of assigning subject fields based on abstracts of the items in question, such as submissions or publications, would be needed to enable this, along with an agreed pan-discipline classification scheme.

## 5.15 Public Engagement ❄
ATTENDEES AT PUBLIC EVENTS

ENDORSED BY: UNITED KINGDOM.

### 5.15.1 Metric definition
Public Engagement ❄ calculates the number of attendees at public events to assess an institution's wider social and cultural impact on their region and nation.

It answers the questions of:
- The number of people that attended public events hosted by a given institution.
- The level of interest or engagement in public events supported by a given institution.
- The extent to which a given institution is engaged in social, community and cultural outreach activities.
- The extent to which a given institution is having an impact on broader social, community and cultural environment, using number of attendees as a proxy.


(a) Number of attendees

(a) Time period
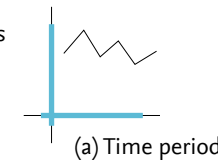
### 5.15.2 Details
Public events are defined as those intended for attendance by the community external to the institution, where measurement of their impact as financial income would not be appropriate. They include knowledge, facility, and cultural awareness events, regardless of whether the events were chargeable or free. Examples of public events are lectures; performance arts such as music, dance, and drama; exhibitions such as those in galleries and museums; and museum education. Open days, student union activity and commercial conferences are excluded.

**5.15.3 Primary data sources**
- Institutional esteem database or Current Research Information System (CRIS system)
- Published annual accounts
- National statutory reports, such as those available from the Higher Education Statistics Agency[164] (HESA) in the UK

**5.15.4 UK application**
Denominator derived from institutional data: institution.

Time period: financial year.

## 5.16 Academic Recognition ✿
ELECTED FELLOWSHIPS OF NATIONAL ACADEMY OF SCIENCES

ENDORSED BY: UNITED KINGDOM.

**5.16.1 Metric definition**
Academic Recognition ✿ calculates the number of elected fellowships to the national academy of sciences.

It answers the question of which institutions are the employers of choice, in that they employ internationally recognised researchers.

(a) Number memberships

(a) Time period
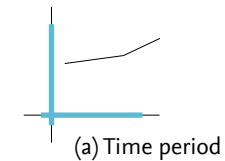
(a) Number memberships per FTE

(a) Time period
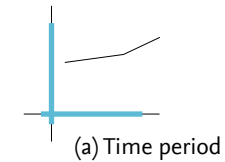
**5.16.2 Details**
Elected fellowships are counted at the university at which the academic is based, and move with the academic. No historic assignment to previous career universities is made.

The number of elected fellowships is counted, not the number of researchers.

Only awards that are open internationally should be counted in this recipe.

The following are not counted:
- Prizes.
- Paid memberships.

**5.16.3 Primary data sources**
- Current Research Information System (CRIS system)
- Membership lists on the national academy websites, combined with a check for the accuracy of the affiliation currency

**5.16.4 UK application**
The following awards are considered to fall under the category of national academies of science:
- Fellow of the Royal Society[165]
- Fellow of the Royal Academy of Engineering[166]
- Fellow of the British Academy[167]

Denominators derived from institutional data: institution.

Time period: calendar year.

**5.16.6 Future opportunities**
It would be valuable to extend the remit of this metric to include a count of competitive early-career fellowships.

## 5.17 Intellectual Property Volume ❄
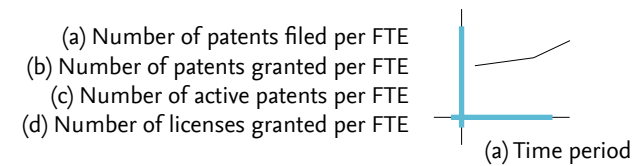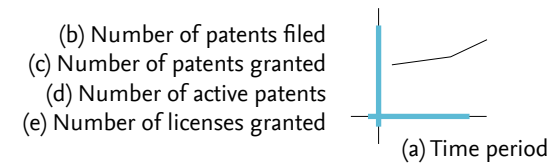
VOLUME OF PATENTS AND LICENSES

ENDORSED BY: UNITED KINGDOM.

**5.17.1 Metric definition**
Intellectual Property Volume ❄ calculates the number of patents that are filed and granted, the number of active patents, and the number of licenses.

It answers the questions of:
- How many genuine innovations an institution produces each year.
- The size of an institution's exploitable portfolio.
- What an institution has exploited commercially that is now being used by industry.

(b) Number of patents filed
(c) Number of patents granted
(d) Number of active patents
(e) Number of licenses granted

(a) Time period

(a) Number of patents filed per FTE
(b) Number of patents granted per FTE
(c) Number of active patents per FTE
(d) Number of licenses granted per FTE

(a) Time period

165 https://royalsociety.org/about-us/fellowship/fellows/

166 http://www.raeng.org.uk/about-us/people-council-committees/the-fellowship/list-of-fellows?fa=a&p=2

167 http://www.britac.ac.uk/fellowship/directory/ord.cfm?letter=A

**5.17.2 Details**

Intellectual Property includes copyrights, trademarks, design rights, trade secrets and patents for the protection of inventions, and licenses granted to private companies allowing them to exploit an institutional invention that is protected by a patent.

The number of patents granted includes all individual patents, and any individual national patents.

The number of patents is the sum of the number of active (currently registered under licence to an external party) and live (registered but yet to be licensed) patents.

The number of licenses granted includes those granted from licence agreements, assignments, exercised option agreements, licences to spin-outs and income-generating Material Transfer Agreements.

**5.17.3 Primary data sources**
- Institutional intellectual property database or Current Research Information System (CRIS system)
- Published annual accounts
- National statutory reports, such as those available from the Higher Education Statistics Agency[168] (HESA) in the UK
- Scopus
- Web of Science
- Google Scholar
- Lexis®[169]
- World Intellectual Property Organization (WIPO)[170]

**5.17.4 UK application**

Denominators derived from institutional data: institution.

Time period: financial year.

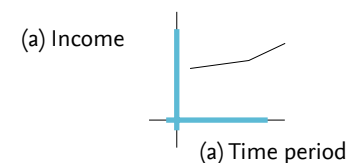## 5.18 Intellectual Property Income ⚙

REVENUE FROM PATENTS AND LICENSES

ENDORSED BY: UNITED KINGDOM.

**5.18.1 Metric definition**

Intellectual Property Income ⚙ calculates the revenue derived from patents and licenses.

It answers the question of how much commercial return an institution is deriving from its interactions with a range of external partners.



(a) Income

(a) Time period

**5.18.2 Details**

Intellectual Property Income is that received by the institution from upfront fees, milestone fees, royalties, and reimbursement of patent costs. Income from design interactions and licensing is also included.

**5.18.3 Primary data sources**
- Institutional accounts system or Current Research Information System (CRIS system)
- Published annual accounts
- National statutory reports, such as those available from the Higher Education Statistics Agency[171] (HESA) in the UK

**5.18.4 UK application**

Denominators derived from institutional data: institution.

Time period: financial year.

Currency: British pounds (GBP).

---

168 www.hesa.ac.uk

169 www.lexis.com

170 www.wipo.int/portal/en

---

171 www.hesa.ac.uk
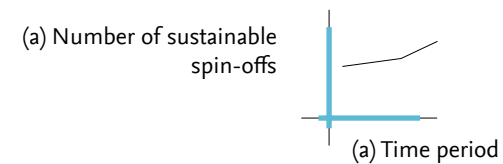
## 5.19 Sustainable Spin-Offs ❄

NUMBER OF SUSTAINABLE SPIN-OFFS

ENDORSED BY: UNITED KINGDOM.

### 5.19.1 Metric definition
This metric calculates the number of sustainable spin-offs.

It answers the question of how many companies that are high quality, and therefore sustainable, an institution has delivered.



(a) Number of sustainable spin-offs

(a) Time period

### 5.19.2 Details
A spin-off is a company that has been set up to exploit intellectual property that originated from within the institution.

The types of spin-off counted in this recipe are those for which the definition is specific and not open to interpretation, and where the institutional data quality upon which the metric is based are relatively high. These are:
- A spin-off with some institutional ownership.
- A spin-off based on institutional intellectual property that is not owned by the institution.

The following types of spin-off are excluded from this recipe, since the definitions are less specific and the data quality is consequently relatively low:
- Staff start-up that has been set up by current institutional staff, or those who were a staff member within the last 2 years, but that are not based on intellectual property from the institution.
- Graduate start-up that has been set up by graduates who are currently members of the institution, or were a member within the last 2 years, regardless of where any intellectual property resides, when there has been formal business support from the institution.
- Undergraduate start up, when there has been formal business support from the institution.

Sustainable spin-offs are active companies that have survived for at least 3 years.

**5.19.3 Primary data sources**
- Institutional intellectual property database or Current Research Information System (CRIS system)
- Published annual accounts
- National statutory reports, such as those available from the Higher Education Statistics Agency[172] (HESA) in the UK

**5.19.4 UK application**
Denominators derived from institutional data: institution.

Time period: financial year.

**5.19.5 Future Opportunities**
It is desirable to agree a more specific definition for the types of spin-offs that are excluded from the count in this recipe, that will support better data collection for more reliable benchmarking results.

There is interest in benchmarking "spin-ins", to indicate the strength of the local innovation system. "Spin-ins" are companies that have not originated from within an institution's research portfolio, but that are based on their premises such as an incubator or science park.

---

172 www.hesa.ac.uk

## 5.20 Spin-Off-Related Finances ❄

FINANCIAL BENEFITS DERIVED FROM ACTIVE SPIN-OFFS

ENDORSED BY: UNITED KINGDOM.

**5.20.1 Metric definition**
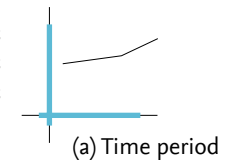This metric calculates the financial benefits derived from an institution's active spin-offs.

It answers the questions of:
- How many jobs an institution is creating from its spin-offs.
- What economic return an institution delivers to its region and / or nation.
- How an institution is helping its companies to grow.
- The quality of an institution's spin-out companies.

(a) Number of FTEs employed by active spin-offs
(b) Turnover from active spin-offs
(c) External investment in active spin-offs

(a) Time period

**5.20.2 Details**
A spin-off is a company that has been set up to exploit intellectual property that originated from within the institution.

The types of spin-off counted in this recipe are those for which the definition is specific and not open to interpretation, and where the institutional data quality upon which the metric is based are relatively high. These are:
- A spin-off with some institutional ownership.
- A spin-off based on institutional intellectual property that is not owned by the institution.

The following types of spin-off are excluded from this recipe, since the definitions are less specific and the data quality is consequently relatively low:
- Staff start-up that has been set up by current institutional staff, or those who were a staff member within the last 2 years, but that are not based on intellectual property from the institution.
- Graduate start-up that has been set up by graduates who are currently members of the institution, or were a member within the last 2 years, regardless of where any intellectual property resides, when there has been formal business support from the institution.
- Undergraduate start up, when there has been formal business support from the institution.

Active spin-offs are those which are currently active, regardless of the number of years that they have existed.

External investment includes all investment from external partners, with the exception of third-stream funds.

These measures may need to be estimated, as is the case for the information returned to the Higher Education Statistics Agency[173] (HESA) in the UK.

**5.20.3 Primary data sources**
- Institutional accounts system or Current Research Information System (CRIS system)
- Published annual accounts
- National statutory reports, such as those available from the Higher Education Statistics Agency (HESA) in the UK

**5.20.4 UK application**
Denominators derived from institutional data: institution.

Time period: financial year.

Currency: British pounds (GBP).

External investment excludes investment from third stream funds from the Higher Education Funding Council for England (HEFCE[174]) and the Department of Business, Innovation and Skills (BIS[175]), which was replaced in 2016 by the Department for Business, Energy & Industrial Strategy (BEIS[176]).

---

173 www.hesa.ac.uk
174 www.hefce.ac.uk
175 www.gov.uk/government/organisations/department-for-business-innovation-skills
176 https://www.gov.uk/government/organisations/department-for-business-energy-and-industrial-strategy

## 5.21 Time to Award of Doctoral Degree ❄

TIME TAKEN TO BE AWARDED A DOCTORAL QUALIFICATION AS A PROPORTION OF EXPECTED COURSE LENGTH

ENDORSED BY: UNITED KINGDOM.

**5.21.1 Metric definition**
Time to Award of Doctoral Degree ❄ calculates the ratio of time for award of a doctoral degree to the expected course length.

It answers the question of the time taken to obtain a doctoral qualification compared to expected course length.



(a) Ratio of time to award of doctoral degree : expected course length

(a) Time period
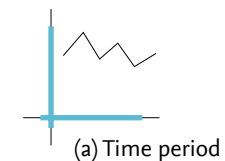
**5.21.2 Details**
- A Time to Award of Doctoral Degree ❄ ratio of 1 means that the expected time was taken to have the degree awarded; above 1 means longer than expected; and below 1 means shorter than expected. However, the expected length of the course excludes the writing up period, and any time that elapses before the degree is examined and approved. It is therefore unlikely that the actual ratio for this metric for any given institution will be 1 or less than one.
- Where there are both part-time and full-time students following a programme of study, the expected length of study should be the normal length applicable for the mode of study (part-time or full-time) of the student in question.
- It is advisable to undertake separate analysis of full-time and part-time students. This is because the amount of administrative time taken to assess and approve a qualification is likely to be the same whether a student studied part-time or full-time, but the assessment and approval time as a proportion of the expected course length (shorter for full-time, longer for part-time students) will differ.

**5.21.3 Primary data sources**
- Institutional systems
- National statutory returns, such as those collected by the Higher Education Statistics Agency (HESA[177]) in the UK

**5.21.4 UK application**
Denominators derived from institutional data:
- HESA Joint Academic Coding System[178] (JACS) code
- Full-time or part-time research students
- Home or overseas research students
- Gender
- Institution

Time period: financial year.

Population: the relevant HESA standard definition should be followed to define the population for the metric calculation. This population is obtained from HESA qualifications, filtered to include only Doctoral awards.

It is noted that JACS will change in the near future[179], in which case the new equivalent field should be used.

The ratio is derived by obtaining the difference between the expected course length, using the HESA fields SPLENGTH[180] and UNITLGTH[181], and the actual time to award, using the HESA fields COMDATE[182] and ENDDATE[183].

**5.21.5 Future Opportunities**
Percentage completion rate would be a helpful additional metric to complement this one.

## 5.22 Destination of Research Student Leavers ❄

RESEARCH STUDENTS CONTINUING IN ACADEMIA FOLLOWING SUCCESSFUL COMPLETION OF COURSE

ENDORSED BY: UNITED KINGDOM.

**5.22.1 Metric definition**
Destination of Research Student Leavers ❄ calculates the proportion of research students who continued to work in the university sector, compared to moving into other career paths.

It answers the question of the proportion of all research students that remained within academia.



(a) Proportion of successful research students continuing within academia

(a) Time period
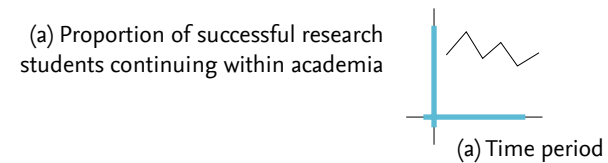
**5.22.2 Details**
This metric is concerned only with successful research students.

**5.22.3 Primary data sources**
- Institutional systems
- National statutory returns, such as those collected by the Higher Education Statistics Agency (HESA[184]) in the UK

---

177 www.hesa.ac.uk

178 www.hesa.co.uk/support/documentation/jacs

179 JACS will be replaced by the Higher Education Costing System (HeCoS) and the Common Aggregation Hierarchy (CAH) for the 2019/20 academic year: www.hesa.ac.uk/innocation/hecos

180 www.hesa.co.uk/collection/c14051/a/splength/

181 www.hesa.co.uk/collection/c14051/a/unitlgth/

182 www.hesa.co.uk/collection/c14051/a/comdate/

183 www.hesa.co.uk/collection/c14051/a/enddate/

184 www.hesa.ac.uk

**5.22.4 UK application**
Denominator derived from institutional data: institution.

Time period: financial year.

Population: the Destination of Leavers from Higher Education (DLHE) Survey[185] organised by the Higher Education Statistics Agency (HESA) in the UK is used to calculate the metric. The DLHE population is based on post-graduate research student respondents who completed one of the two annual DLHE survey collections. The POPDLHE field, filtered to include the XDLEV501 code 1, forms the basis of the population. The DLHE highest qualification obtained (XOBTND01) field is used to identify masters and doctoral research student leavers (code L00 for masters research, and D00 and D01 for doctoral research). Standard Occupational Classification[186] (SOC2010) codes for researchers and post-doctoral researchers (21191, 24269 and 31110) and lecturers/academic staff (23110) in combination with the Standard Industrial Classification[187] (SIC2007) 2-digit SIC code 85* are used to determine the proportion of the population who have remained in academia following graduation. For the purposes of this metric, all other respondents are considered to have taken another career path.

**5.22.5 Future Opportunities**
It would be of interest to extend this metric to be able to, for example, benchmark based on the leavers' destinations by funder type, leavers entering industry and moving overseas.

---

185 www.hesa.co.uk/support/definitions/destinations

186 www.hesa.co.uk/support/documentation/occupational/soc2010

187 www.hesa.co.uk/support/documentation.industrial.sic2007

# 6. Snowball Metrics cards

An example of a Snowball Metrics card is shown below. Metrics cards for all Snowball Metrics can be downloaded for use from the Snowball Metrics website[188].

You are free to use these cards to help facilitate discussions around Snowball Metrics, for training sessions and other purposes.



---
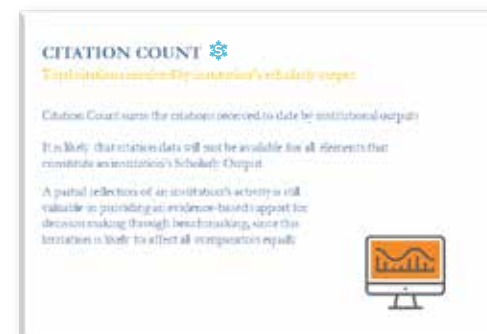
188 www.snowballmetrics.com/metrics/

# 7. Index of denominators and metrics

## Denominators

# 7. Index of denominators and metrics

## Metrics

## 8. Notes