Snowball Metrics

Response to the call for evidence to the independent review of the role of metrics in research assessment

Dr John T. Green, Queens' College, Cambridge, Chair of the Snowball Metrics Steering Committee
jtg11@cam.ac.uk +44 7725 159862

Representatives of the Snowball Metrics Programme Partners would be interested in participating in a workshop/event to discuss the use of metrics in research assessment and management

---

This response has the following sections:

1. **Executive Summary**
2. **Advocacy for Snowball Metrics**
3. **Background to, and status of, the Snowball Metrics programme**
4. **Adoption and globalisation of Snowball Metrics**
5. **Specific responses to HEFCE questions**

1. Executive Summary

Whilst recognising that metrics are only a part, albeit an important part, of the evaluation landscape, Snowball Metrics offer a robust framework for measuring research performance and related data exchange and analysis. Snowball Metrics should be endorsed and further developed as the UK standard to align how key metrics are defined across HEIs, HEFCE, HESA, Research Councils and other stakeholders. Snowball Metrics promote efficiencies in data collection and analysis, and enable international comparisons not only on publications but also on many other metrics.

Eight high-profile UK universities[1] started working together in 2010 to address shared challenges

- Universities need robust metrics to create evidence-based strategies and to enable us to: identify our strengths; decide where we want to focus, through investing and divesting; identify with whom we collaborate; facilitate collaboration; understand our competitors and identify peer groups; benchmark against our collaborators and competitors; connect our researchers; and encourage mobility.
- It can be very difficult to find quality management information which can be used with confidence.
- There is no consistent approach to information and measurement between institutions, funders and government bodies, resulting in duplicated effort and inefficiency across the sector, including HEFCE & BIS.

Our major aims

- To understand the aspects of research and related activity we want to measure in order to manage our institutions effectively.
- To do this by defining standardised metrics for measuring in unambiguous and rigorous ways (across the whole range of research activity: funding, collaboration, commercialisation) so as to ensure that we are 'comparing apples with apples'.
- To achieve real-time snapshots of data at best, and annual snapshots at worst.

Our fundamental principles

The Snowball Metrics Programme is based on three fundamental principles which we believe are vital to the successful adoption of metrics:

- the development of metrics must be sector led;
- the methodologies, or "recipes", developed are open source and supplier agnostic;
- the community is the guardian of the metrics.

[1] 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

Our progress

- Over four years we have made huge progress in creating robust definitions for 24 metrics accepted by the eight partners and gradually adopted globally. These methodologies are openly available in the Snowball Metrics recipe book, at www.snowballmetrics.com/metrics.
- We have had keen interest from UK university representative groups and from the vast majority of UK research intensive universities.
- We have gained worldwide traction and endorsement (see 5.4 below).
- We are working closely with HESA, EuroCRIS[2] and CASRAI[3] to build on their work and expertise as organisations that have successfully developed standard definitions, data structures and dictionaries which work across national boundaries.
- We are currently starting to use natural language processing methodologies to tackle one of the hardest and most pressing problems of finding a thematic taxonomy that all of our inputs, outputs and outcomes can be mapped to.

Our plea

- **It would be counter-productive to reinvent the wheel rather than build upon the huge experience and knowledge base that has been established within the Snowball Metrics Programme: a significant number of high profile universities have been working on this for four years not only in the UK, but also in the US, and Australia / New Zealand.**
- **As a result, Snowball Metrics are firmly established in the UK and are being accepted internationally (5.4 below).**
- **If the sector chooses to develop the use of metrics, then we urge the sector to build upon the Snowball Metrics Programme.**
- **Metrics need an appropriate neutral body to manage them which ensures that they are owned by the sector but which also ensures maintenance and upkeep; this is little different from the REF, much of which is felt to be owned by the sector through UoA panels, but which needs third parties – HEFCE and subcontractors – to manage the process. We suggest that consideration be given as to which neutral body is most appropriate, perhaps HESA, HEFCE or BIS.**

2. Advocacy for Snowball Metrics

The following have contributed to this response and endorse the following statement:

We support the approach and vision of the Snowball Metrics Programme. If it is decided that the role of metrics be considered further then we very much hope that HEFCE will take the opportunity to consider Snowball Metrics as an existing approach which has been robustly created and which is gaining significant traction both within the UK and globally.

University of Oxford Glenn Swafford, Director of Research Services
University of Cambridge Malcolm Edwards, Head of Planning and Resource Allocation
University College London Jacob Sweiry, Principal Research Facilitator, Office of the Vice Provost – Research
Imperial College London Ian McArdle, Research Systems and Information Manager
University of Bristol David Langley, Director, Research and Enterprise Development
University of Leeds Jennifer Johnson, Head of Performance, Governance & Operations
Queen's University Belfast Scott Rutherford, Director, Research and Enterprise
University of St Andrews Anna Clements, Head of Research Data and Information Services
Higher Education Statistics Agency Alison Allden, Chief Executive

[2] euroCRIS, a not-for-profit organisation interested in current research information systems and their interoperability: www.eurocris.org
[3] CASRAI, Consortia Advancing Standards in Research Administration, is a non-profit, standards development organisation: www.casrai.org

3. Background to, and status of, the Snowball Metrics programme

Snowball Metrics began as a response to a HEFCE/JISC report in 2010 which highlighted challenges around data and metrics that are faced by universities[4]. It is an ongoing collaborative partnership between eight UK universities and Elsevier:

- 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

These universities together account for nearly 40% of competitive funding awarded by the UK's Research Councils, and about 40% of UK-authored articles and of UK citations. The ambition is that the conclusions and approaches endorsed by this core will gain traction throughout the international higher education sector through a "snowball effect", hence the name Snowball Metrics; indeed this ambition is coming to fruition.

The aim of Snowball Metrics is to become the international standard that is endorsed by research-intensive universities that enable them to understand their strengths and weaknesses, so that they can build and monitor effective strategies (e.g. in which areas to invest, in which to divest, which people to promote from where).

Snowball Metrics enable informed, evidence-based decision-making by agreeing a single method to calculate metrics that will provide input to institutional strategies by ensuring that apples are compared with apples. These metrics are based on all the data sources available to us, including institutional data sources, as well as third party and commercially available sources**.**

Snowball Metrics do not depend on a particular data source or supplier, and are owned by the higher education sector. **Snowball Metrics is, at its heart, a bottom-up initiative**.

The process through which the recipes have been developed is based on methodologies which ensure **robust and unambiguous definitions** so that the metrics they describe enable the confident comparison of apples with apples. The resulting benchmarks can be trusted as reliable information to help establish and monitor institutional strategies.

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.

Our ambition is that Snowball Metrics become internationally recognised by their snowflake kitemark and are endorsed by universities as a standard to illuminate the strengths and weaknesses of universities. We urge institutions, funders, agencies, and suppliers of research information to adopt Snowball Metrics. A single method of requesting and consuming information will drive enormous efficiencies in all sectors of higher education.

We aim to define Snowball Metrics throughout the entire landscape of research activities, and so far we have 24 recipes available for free to the sector. In addition to agreeing the metrics methodologies themselves, a set of **denominators** is needed to enable measurement by theme, by institution, by discipline, etc. These denominators enable: **slicing and dicing** the Snowball Metrics at levels that are more granular than an entire institution in order to understand strengths within a discipline, since inputs such as funding, or outputs such as articles, vary considerably (between disciplines, for example); and **normalising for size** between institutions, so that it is not always the case that bigger institutions appear to perform better. The following is the landscape of Snowball Metrics and of denominators (from Figure 2[5]).

[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

[5] http://www.snowballmetrics/metrics

	Research Inputs	*Research Process*	*Research Outputs and Outcomes*
Research	**Applications Volume** **Awards Volume**	**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 **Collaboration** • Collaboration • Collaboration Impact • Academic-Corporate Collaboration • Academic-Corporate Collaboration Impact **Societal impact** • Altmetrics • Public Engagement
Enterprise Activities/ Economic Development	• Academic-Industry Leverage • Business Consultancy Activities	• Contract Research Volume	• Intellectual Property Volume • Intellectual Property Income • Sustainable Spin-Offs • Spin-Off-Related Finances
Post-Graduate Education			

Denominators	**(Number of) People**	**Organisations**	**Themes / Schemes**
• "Slice and dice" • Normalize for size	• Researcher, authors • Principal / co-investigators • Academic staff by category • Research assistants • PGR Students • UG / PGT Students • Post-doctoral staff • Support staff	• Institution • Faculty / department • Cost Centre • Groups / clusters • Funders by type • Centres / institutes	• Standard grants • Strategic initiatives (Calls) • Grand Challenges • Subject areas • Keywords

The first edition of the Snowball Metrics Recipe Book[6] was published in 2012 with 10 recipes including volume of research grant funding submitted to external funding bodies, volumes awarded, volume of research income spent market share, bibliometrics and metrics to measure international collaboration; the second edition[7] was published in 2014 with a further 14 metric recipes, mostly in the areas of collaboration, enterprise and impact. The third phase, on which we have now embarked, will focus on metrics in post-graduate education and collaboration, as well as a thematic subject classification.

[6] http://www.snowballmetrics.com/wp-content/uploads/snowball-metrics-recipe-book-upd.pdf
[7] www.snowballmetrics.com/metrics

4. Adoption and globalisation of Snowball Metrics

The UK Steering Group developed a pilot tool, the Snowball Metrics Lab that Elsevier built for working on the data available to members of this Group, to support the feasibility testing of the recipes before they were published and shared with the sector. This pilot gave rise to the concept of the **Snowball Metrics Exchange.** Elsevier have committed to building an API during 2014 which is a free "broker service" for the exchange of Snowball Metrics between peer institutions who agree that they would like to share information with each other:

- any institution using Snowball Metrics can become a member of the Snowball Metrics Exchange;
- the institutional members will be responsible for generating their Snowball Metrics according to the recipes, whether they are 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; 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;
- **the data underlying the metrics will never be exchanged, only the metrics, and will remain behind the institutions' firewalls.**

Interest in metrics has intensified in other geographies. In the US, a report[8] in 2013 raised similar issues to those recognised in the UK which then gave rise to the formation of a **US Snowball Working Group** comprising seven large US universities - a mix of public and private, with and without medical schools, whose objectives are

- **endorsing** all, or as many as possible, Snowball Metrics to drive the move towards global standards;
- **enhancing** existing Snowball Metrics with national data and intelligence;
- **enabling** global benchmarking using national data by understanding how to **map** national denominators for cross-country compatibility;
- **developing** new metrics for the global initiative if there are gaps from the national perspective.

This group has worked with members of the UK Steering Group and has made good progress against its aims including refining some of the UK's early recipes to make them both globally and nationally workable.

A **Working Group** has been formed in **Australia / New Zealand**, composed of 8 universities. **Portugal** is using Snowball Metrics in their upcoming national assessment[9]. The **RU11 group of Japanese** research-intensive universities has established a Metrics Working Group and is working with the UK Steering Group.

The conclusion is that Snowball Metrics are firmly established in the UK and are being accepted internationally. It would be counter-productive to reinvent the wheel rather than build upon the huge experience and knowledge base that has been established within the Snowball Metrics Programme.

.

5. Specific responses to HEFCE questions

5.1 Identifying useful metrics for research assessment

- ***What empirical evidence (qualitative or quantitative) is needed for the evaluation of research, research outputs and career decisions?***

  Metrics complement peer review and expert opinion; the ideal situation is to have information from all three types of input. If intelligence from these complementary approaches "triangulates", i.e. gives a consistent picture, then this increases the confidence in conclusions. Inconsistent views might suggest that further investigation is needed. It is also advisable to "triangulate" within the metrics corner of the triangle, and this is one reason that Snowball Metrics aim to agree on a broad set of metrics. Another reason of course is the broad diversity of questions that they could be used to help address. Snowball Metrics provide a balanced scorecard of metrics from which a selection can be made.

[8] "Evidence-based decision making in academic research: The "Snowball" effect", published in the Academic Executive Brief, 2013. http://academicexecutives.elsevier.com/articles/evidence-based-decision-making-academic-research-snowball-effect

[9] Cristiana Leandro (Cristiana.leandro@fct.pt) is the executive coordinator of the Scientific Council for Exact Sciences and Engineering within FCT https://www.fct.pt/fct.phtml.en

- ***What metric indicators are currently useful for the assessment of research outputs, research impacts and* research environments? *What new metrics, not readily available currently, might be useful in the future?***
  See Fig.2 in the Snowball Metrics Recipe Book (2nd edition[10], reproduced above) for the landscape of metrics which the Snowball Metrics Programme has defined as useful, both those in place now and whose which we intend to cover in the future. See the Snowball Metrics Recipe Books for the 24 recipes already defined across all areas of research activity (inputs, throughputs and outputs).
- ***Are there aspects of metrics that could be applied to research from different disciplines?***
  Some Snowball Metrics will be more applicable to some disciplines than others and at different stages of the research process (inputs, processes, outputs and outcomes). Key is that metrics complement peer review and expert opinion in varying degrees of balance, depending on the question, discipline and subject as well as on the specific expertise of the questioner.
  Snowball Metrics are categorised / normalised by discipline so apples are compared with apples. Given a requisite level of data coverage and quality, all metrics could be applied across disciplines. So in Arts and Humanities (A&H), whilst grant income is generally low, peers are still compared to peers within the discipline; coverage for books, chapters etc is generally poor in major publication databases and therefore citations are less relevant to A&H disciplines. In some fields, esteem becomes more important; for example, there are geographers, philosophers and so on who are FBA yet would not stand out when measured by grant income or h-index. In STEM, some disciplines have high output volumes – physics, genetics and so on – but again Snowball Metrics normalise by discipline to ensure apples are compared with apples. The point of Snowball Metrics is the rounded picture of performance they give. When talking about quality the best researchers will demonstrate performance across the board, the only caveat perhaps being impact where measures are less defined and quantum-based.
- ***What are the implications of the disciplinary differences in practices and norms of research culture for the use of metrics?***
  The concept of "denominator" as exploited by Snowball Metrics is critical and applies to all Snowball Metrics recipes. This enables measurement by theme, by discipline, by institution, by department etc. These denominators enable: **slicing and dicing** the Snowball Metrics at levels that are more granular than an entire institution in order to understand strengths within a discipline, for instance; and **normalising** for different sizes of disciplines within institutions.
- ***What are the best sources for bibliometric data? What evidence supports the reliability of these sources?***
  Snowball Metrics are agnostic of data source and can be generated using any of (e.g.) Web of Science, Scopus, and Google Analytics; when comparing (benchmarking) Snowball Metrics calculated by different universities, the data sources used in the recipes must be the same (apples to apples). Snowball Metrics have no preference or dependency on tools from third party suppliers to calculate the metrics from the recipes; indeed they can be calculated independently by bespoke institutional systems.
- ***What evidence supports the use of metrics as good indicators of research quality?***
  Much of UK funding (QR, RCUK, TSB, etc.) is allocated through some mixture of metrics-based quality assessment of performance combined with peer-review. There is increasing reliance on metrics particularly for custom analytics reports such as those regularly produced for governments (see for example International Comparative Performance of the UK Research Base – 2013 a report for the UK's Department of Business, Innovation and Skills)[11].

[10] www.snowballmetrics.com/metrics

[11] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/263729/bis-13-1297-international-comparative-performance-of-the-UK-research-base-2013.pdf

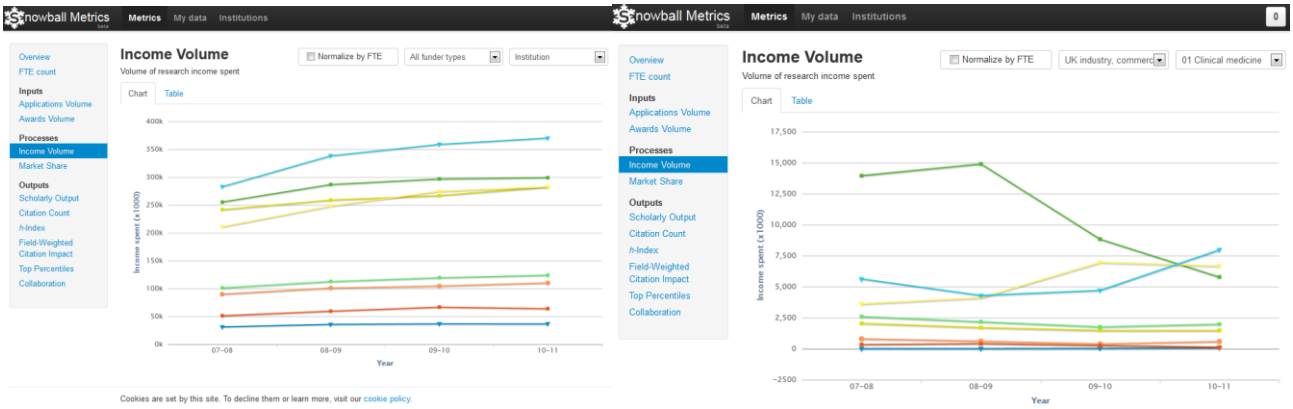
- ***Is there evidence for the move to more open access to the research literature to enable new metrics to be used or enhance the usefulness of existing metrics?***

  The Snowball Metrics Programme takes very seriously the need to protect the universities' preferred approach from the impositions of external stakeholders (governments, suppliers, funders etc). The UK Steering Group considers and develops new metrics and does this through discussion amongst experts (including Elsevier as a provider). It is through that consensus we decide what is feasible and reliable rather than by imposition from government bodies or funders with perhaps narrower objectives. (The most recently published Snowball Metrics include, for example, the recipe for Altmetrics).

5.2 How should metrics be used in research assessment?

- ***What examples are there of the use of metrics in research assessment?***

  There are numerous examples of the use of metrics by policy makers, funders, universities and increasingly industry. When the Snowball Metrics programme partners first saw the pilot tool enabling them to compare metrics amongst themselves, their value was compelling, particularly when they were able to slice and dice (e.g. award data by funder type, by discipline, by researcher etc):



- ***To what extent is it possible to use metrics to capture the quality and significance of research?***

  Metrics can indicate quality and impact to varying degrees but must always be used alongside peer review and expert opinion. The balance of these inputs into a decision is most likely different by discipline and the question being asked.

- ***Are there disciplines in which metrics could usefully play a greater or lesser role? What evidence is there to support or refute this?***

  The pragmatic approach of Snowball Metrics means that there is currently an imbalance between the Snowball Metrics that can be applied to the social sciences and arts and humanities compared with those that are perhaps most relevant to STEM. While metrics such as Applications Volume[12], Awards Volume[13], Income Volume[14], and Market Share[15] are equally useful across all fields, when the disciplinary denominator is used, metrics such as Citation Count[16] and Collaboration[17] may be less valuable in the non-STEM areas. It is the aim of Snowball Metrics to define recipes that support benchmarking across all disciplines, as well as across all university activities: this is illustrated in the recipe book by the clarifications of definition of Scholarly Output[18], and by the inclusion of Altmetrics[19] and Public Engagement[20].

---

[12] Applications Volume calculates the number and price, or amount applied for, of research grant applications that are submitted to external funding bodies.

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

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

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

[16] Citation Count sums the citations received to date by institutional outputs.

[17] Collaboration calculates the number and percentage of outputs that have national or international co-authorship.

[18] Scholarly Output counts the number of institutional outputs of any type.

[19] Altmetrics counts the number of online events that have been stimulated by an institution's output.

- ***How does the level at which metrics are calculated (nation, institution, research unit, journal, individual) impact on their usefulness and robustness?***
  Critical to the definition of Snowball Metrics is the consistent and generically-used concept of denominator (see Figure 2 [21] to demonstrate the approach).

5.3 'Gaming' and strategic use of metrics:

- ***What evidence exists around the strategic behaviour of researchers, research managers and publishers responding to specific metrics?***
  Whilst there is evidence that researchers and journal editors occasionally game metrics (e.g. through self-citation), Snowball Metrics are unlikely to distort the research process in unanticipated ways through encouraging too much focus on a particular activity. Snowball Metrics offers a balanced scorecard, rather than a focus on one or only a few metrics, and so it remains the decision of the researchers or institutions where they should focus their efforts. The purpose of Snowball Metrics is to understand institutional strengths and weaknesses, so that this intelligence can be used to inform university strategies.
  The Snowball Metrics Programme has developed a mixture of performance measures, delivering what is feasible but at the same time not simply measuring what is easy. Often the easy measures get used (usually income, or publication or citation counts) which can indeed drive certain behaviours; sometimes too broad a set of measures are developed in an effort to please all and thus the important metrics can get lost (e.g. HEBCIS measures a raft of indicators). Snowball has aimed at a balanced scorecard and prioritisation.
- ***Has strategic behaviour invalidated the use of metrics and/or led to unacceptable effects?***
  It is critical that the universities who are exchanging Snowball Metrics can trust the underlying data, even though they cannot see the actual data itself, as in the case of metrics generated from other institutions' data. The driver behind Snowball Metrics is internal strategy, and not for showcasing or ranking, so there there is little if any motivation to "game" the metrics, because there is no gain for an institution in concealing its standing amongst its peers from itself.
- ***What are the risks that some groups within the academic community might be disproportionately disadvantaged by the use of metrics for research assessment and management?***
  Disciplines with poor coverage in commercial databases are disadvantaged. But the more that metrics are recognised as robustly and rigorously defined then then the more will data be harvested to drive acceptance. In fact the acceptance of Snowball Metrics is emphasising the data and the format in which it is collected and thereby improving coverage through a reinforcement cycle.
- ***What can be done to minimise 'gaming' and ensure the use of metrics is as objective and fit-for-purpose as possible?***
  Robustly defined metrics inherently minimise gaming – provided, as with Snowball Metrics, the recipes are unambiguous and the data sources specified. Metrics should be used in conjunction with one another (and other evaluative methods) so that gaming one would have no effect on the picture which emerges (imagine a jigsaw – the picture emerges even before every piece is in place – and still becomes clear even if the dog has chewed one or two pieces). Ensuring that metrics are owned by academics, rather than imposed by administrators, will encourage a culture amongst researchers of not wanting to game their behaviours.

[20] Public Engagement calculates the number of attendees at public events.
[21] http://www.snowballmetrics.com/metrics

5.4 International perspective (relevant evidence and examples from outside of the UK)

Institutions compete globally for resources, share equipment and collaborate internationally. Therefore there is a need for universal metrics which are comparable across national boundaries (whilst recognising the need for national flavours, as have been developed amongst the Snowball Metrics).

As mentioned above there is an increasing interest in Snowball Metrics from outside the UK evidenced by:

- the US Snowball Metrics Working Group;
- an Australia/New Zealand Snowball Metrics Working Group;
- progress by Tsukuba University in leading to create a Snowball Metrics Working Group amongst the RU11 Group of Japanese research-intensive universities;
- enthusiasm from Asia Pacific universities consortium to engage
- numerous invitations to present and follow-through with colleagues from governments (including the European Commission), and universities in the EU Russia, Poland, Turkey and Asia-Pacific countries.

It is worth noting that the US Snowball Metrics Working Group was established in late 2013. Although there was initial scepticism that it would not be possible to translate the recipes from the UK into a US language (because of different data definitions and sources) after six months each of the original Snowball Metrics which the Group has to date worked upon has been endorsed by the US (with improvements made to some of the definitions and using US data sources mapped to UK sources).

The second edition of the Snowball Metrics Recipe Book[22] contains quotations giving support and referencing their international importance, including from:

Professor Ian Walmsley, Pro-Vice-Chancellor (Research, Academic Services and University Collections), University of Oxford
David W. Richardson, Associate Vice Chancellor for Research, Director of Sponsored Programs, University of Illinois at Urbana-Champaign, US
Professor Jun Ikeda, Chief Advisor to the President, University of Tsukuba, Japan
Euan Adie, founder of Altmetric[23], United Kingdom
Mark Connelly, Director, Research Fish[24], United Kingdom

Many have written in support of the Snowball Programme, typical of which is the following[25]

Christopher Tremewan Secretary General, **Association of Pacific Rim Universities (APRU)**[26]

"For universities and countries to benchmark their performance, to make the right decisions about both collaborating and competing, and to overcome the deleterious effects of spurious international rankings, Snowball Metrics is an initiative which could be the basis of a cost-effective, trustworthy global benchmarking regime. This would be of great value to the rising institutions of the Asia-Pacific region."

[22] www.snowballmetrics.com/metrics
[23] www.altmetric.com a supplier who has committed to using Snowball Metrics in their next version of software
[24] https://www.researchfish.com/
[25] Others include: Professor Jan Fazlagić, ProRector, Vistula University, Poland; Professor Vladimir Kruzhaev, Vice Rector, Ural Federal University; Professor Victor Soifer, President, Samara State Aerospace University, Russia; Professor Igor Osipov, School of Business, University of Alberta, Canada; Dr Yukihito Morimoto & Dr Masayo Shindo, University of Tsukuba, Japan
[26] http://apru.org/about/welcome