

A new metric for measuring journal impact – Usage Factor

Peter Shepherd
Director, COUNTER

STM

7 December 2012

Usage Factor: the context

- **A growing body of reliable journal usage statistics**
 - The burgeoning availability of reliable usage data for online journals has opened the door to usage-based measures of impact, value and status.
 - Since 2002 COUNTER has provided a global standard for usage statistics
 - Over 15,000 full text journals now covered by COUNTER standards
 - COUNTER developing a new Code of Practice –PIRUS- that will also cover repositories
- **A complement to altmetrics and citation-based measures of impact**
 - Impact Factors, based on citation data, are generally accepted as a valid measure of the impact and status of scholarly journals
 - Do not cover all fields of scholarship
 - Do not reflect the true impact of more practitioner-oriented journals
 - Altmetrics, take into account novel forms of scholarly communication
 - Repositories, archives, blogs, social media
 - More immediate measure of impact than citations
 - Scholar and item based rather than journal-based
 - What do the numbers mean?

Who will benefit from the Usage Factor?

Four major groups will benefit from the introduction of Usage Factors:

- **Authors**, especially those in practitioner-oriented fields, where citation-based measures understate the impact of journals, as well as those in areas outside the core STM fields of pure research, where coverage of journals by citation-based measures is weak.
- **Publishers**, especially those with large numbers of journals outside of the core STM research areas, where there is no reliable, universal measure of journal impact, because citation-based measures are either inadequate or non-existent for these fields
- **Librarians**, when deciding on new journal acquisitions, have no reliable, global measures of journal impact for fields outside the core STM research fields. They would use usage-based measures to help them prioritise journals to be added to their collections.
- **Research Funding Agencies**, who are seeking a wider range of credible, consistent quantitative measures of the value and impact of the outputs of the research that they fund.

Usage Factor Project

- aims and objectives

The **overall aim** of this project was to explore how online journal usage statistics might form the basis of a new measure of journal impact and quality, the Usage Factor for journals.

Specific objectives were to answer the following questions:

- Will Usage Factor be a statistically meaningful measure?
- Will Usage Factor be accepted by researchers, publishers, librarians and research institutions?
- Will Usage Factor be statistically credible and robust?
- Is there an organizational and economic model for its implementation that would cost-effective and be acceptable to the major stakeholder groups.

The project is being carried out in **three Stages**:

- Stage 1 (2007-2008): market research
- Stage 2 (2009-2011): modelling and analysis
- Stage 3 (2011-2012): further tests based on draft Code of Practice

Usage Factor Stage 1

- market research

Interviews with 29 key authors/editors, librarians and publishers; web-based survey of 155 librarians; web-based survey of 1400 authors.

Key findings:

- the majority of publishers were supportive of the UF concept, appeared to be willing, in principle, to participate in the calculation and publication of UFs, and were prepared to see their journals ranked according to UF
- the great majority of authors in all fields of academic research would welcome a new, usage-based measure of the value of journals
- UF, were it available, would be a highly ranked factor by librarians, not only in the evaluation of journals for potential purchase, but also in the evaluation of journals for retention or cancellation
- COUNTER was on the whole trusted by librarians and publishers and was seen as having a role in the development and maintenance of the UF

Usage Factor Stage 2

- modelling and analysis

- Real journal usage data analysed by John Cox Associates, Frontline GMS and CIBER
- Participating publishers: American Chemical Society, Emerald, IOP, Nature Publishing Group, OUP, Sage, Springer
- 326 journals
 - 38 Engineering
 - 32 Physical Sciences
 - 119 Social Sciences
 - *29 Business and Management*
 - 35 Humanities
 - 102 Medicine and Life Sciences
 - *57 Clinical Medicine*
- c.150,000 articles

Recommendations: the metric

- Usage Factors should be calculated using the median rather than the arithmetic mean
- A range of Usage Factors should ideally be published for each journal: a comprehensive UF (all items, all countable versions) plus supplementary factors for selected items
- Usage Factors should be published as integers with no decimal places
- Usage Factors should be published with appropriate confidence levels around the average to guide their interpretation
- The Usage Factor should be calculated initially on the basis of a maximum usage time window of 24 months.
- The Usage Factor is not directly comparable across subject groups and should therefore be published and interpreted only within appropriate subject groupings.
- The Usage Factor should be calculated using a publication window of 2 years

Recommendations: the metric

- Small journals and titles with less than 100 downloads per item may be unsuitable candidates for Journal Usage Factors: these are likely to be inaccurate and easily gamed
- The Usage Factor provides very different information from the citation Impact Factor and this fact should be emphasised in public communications.
- Further work is needed on Usage Factor gaming and on developing robust forensic techniques for its detection
- Further work is needed to broaden the scope of the project over time to include other usage-based metrics

The Journal Usage Factor

- the calculation

Publishers will be able to generate Usage Factors using the Code of Practice, but will have to be independently audited for their Usage Factors to be listed in the Usage Factor Central Registry

- The Journal Usage Factor 2009/2010: all items

The median number of successful requests during 2009 and 2010 for countable items published in the journal during 2009/2010

- Different items types have different impacts in different fields

- The Journal Usage Factor 2009/2010: full-text articles

The median number of successful requests during 2009/2010 for full-text articles published in the journal during 2009/2010

- Challenges:

- Consolidation of usage data from different sources
- Consistent item type definitions
- Economic/Organizational model to support the Central Registry

Recommendations: infrastructure

- Development of systems to automate the extraction and collation of data needed for UF calculation is essential if calculation of this metric is to become routine
- Development of an agreed standard for content item types, to which journal specific item types would be mapped, is desirable as it would allow for greater sophistication in UF calculation
- Development or adoption of a simple subject taxonomy to which journal titles would be assigned by their publishers
- Publishers should adopt standard “article version” definitions based on ALPSP/NISO recommendations

Stage 3: objectives

- Publication of a draft Code of Practice for the Usage Factor, which provides the basis for publishers to calculate and report Usage Factors
- Further testing of the recommended methodology for calculating Usage Factors for journals
- Investigation of an appropriate, resilient subject classification scheme for the classification of journals
- Exploration of the options for an infrastructure to support the sustainable implementation of Usage Factor
- Development of an independent audit process to monitor the calculation and reporting of Usage Factors by publishers (an extension of the COUNTER audit which already takes place)
- Investigate the feasibility of applying the Usage Factor concept to other categories of publication in addition to journals

Stage 3 :

the draft Code of Practice

- The Code of Practice will be consistent with COUNTER and will provide:
 - A list of Definitions and other terms that are relevant to Usage Factor
 - A methodology for the calculation of Usage Factor as a median value, including specifications for the metadata to be recorded, the content types and article versions whose usage may be counted, as well as the Publication Period and Usage Period to be used.
 - NISO/ALPSP best practice on journal article versions adopted and ‘countable article versions’ specified
 - Specifications for the reporting of the Usage Factor
 - Data processing rules to ensure that Usage Factors are credible, consistent and compatible, including protocols for identifying and dealing with attempts to game the Usage Factor
 - Specifications for the independent auditing of Usage Factors
 - A description of the role of the Central Registry for Usage Factors in the consolidation of usage data and in the publication of Usage Factors
- The draft Code of Practice was published in March 2012 for public comment.
- Publishers are invited to prepare UFs using the draft Code of Practice

Stage 3: further tests of methodology and process

- Issues tested
 - 12-month and 24-month Usage Factor periods
 - Stability of UF for low usage journals
 - Stability of UFs over time
 - Additional gaming scenarios
 - Process for data collection and consolidation
 - Format
 - Metadata
 - Central Registry model
- Scope of tests
 - 27 subject fields
 - 11 publishers + 2 aggregators (ACS, AMA, AIP, BMA, Emerald, IOP, Nature, Sage, Springer, Wiley, Wolters Kluwer + EBSCO, Social Science Research Network)
 - Ca 200 journals

Stage 3: subject classification scheme

- **Criteria**
 - Must cover all the major fields of scholarship
 - Must have an appropriate level of granularity, with 300-400 subject/sub-subject categories
 - Must be international, with no particular geographic bias
 - Must be readily accessible, without barriers, to the organizations wishing to use it
 - Must be maintained and updated by an independent third party that is trusted by scholars, publishers and librarians
- **Candidate subject classification schemes**
 - Several schemes examined
 - Ringgold Types Scheme, developed for classification of institutions but applicable to journals, looks very promising
- **Implementation**
 - Publishers will be invited to allocated their publications to subject fields within the approved scheme
 - Usage Factor International Advisory Board will review the publisher allocation and modify where necessary
 - Publishers will have the right to appeal decisions of the International Advisory Board

Stage 3: infrastructure

The Central Registry

The Central Registry is the key part of the infrastructure that will support the implementation of Usage Factors on a sustainable, global basis. It will have two main functions:

- To collect and process usage data at the individual item level from publishers and other sources in order to derive Consolidated Usage Factors per publication.
- To provide a central, open source Registry of valid Publisher Usage Factors and Global Usage Factors

To test these functions an important part of the Stage 3 of the project is to set up a Central Registry Demonstrator, focussing on journal articles, which will have the following capabilities:

- Collect usage data from publishers, aggregators and other sources, using the format specified
- Automatically consolidate, from the different sources, the usage data for each article
- Calculate the Usage Factor for each journal, as the Median value specified in Section 3 of the Code of Practice for Usage Factors
- Report the Usage Factors for each journal, as specified in Section 3.2.1 of the Code of Practice for Usage Factors
- Calculate Usage Factors for a number of consecutive Usage Periods to test the stability of the measure over time.

Stage 3: contractor and timetable

The contractor for the collection and consolidation of usage data from publishers/aggregators and for the statistical analysis of the data is CIBER

Timetable:

- November 2012 : submission of data by publishers to CIBER in the prescribed format
- December 2012: initial testing of data and further clarification of issues
- December 2012/January 2013: main analysis phase
- February 2013: interim report and discussion of findings
- April 2013: final report

Beyond Stage 3.....

Once Stage 3 of the Usage Factor project is completed, assuming the results are satisfactory, it is envisaged that Usage Factor will go to full implementation, which will include:

- Publication of the definitive Release 1 of the Code of Practice for Usage Factor
- Setting up the governance structure for Usage Factor
- Setting up the Central Registry and other aspects of the required Usage Factor infrastructure
 - Technical model
 - Business model
- Invitation to publishers to participate in the recording and reporting of Usage Factors for their journals

Extending the scope of Usage Factor beyond journals

- Article/researcher level Usage Factors
 - Usage Factor data being collected at the article level
 - ORCID identifier will facilitate the collection of usage data for an individual researcher
- Usage Factors for other publications
 - Online books
 - Online databases (text, image, video, etc)

Usage Factor : Stage 3 organization

Project Co-Chairs:

- Jayne Marks, Wolters Kluwer, USA
- Hazel Woodward, Cranfield University, UK

International Advisory Board Members

- Mayur Amin, Elsevier, UK
- Kim Armstrong, CIC Center for Library Initiatives, USA
- Peter Ashman, BMJ Group, UK
- Terry Bucknell, University of Liverpool, UK
- Ian Craig, Wiley, UK
- Joanna Cross, Taylor & Francis, UK
- David Hoole, Nature Publishing Group, UK
- Tim Jewell, University of Washington, USA
- Jack Ochs, ACS Publications, USA
- Tony O'Rourke, IOP Publishing, UK
- Clive Parry, Sage Publications, UK
- Jason Price, Claremont College, USA
- Ian Rowlands, CIBER, UK
- Bill Russell, Emerald, UK
- Ian Russell, Oxford University Press, UK
- John Sack, HighWire Press, USA
- David Sommer, COUNTER, UK
- Harald Wirsching, Springer, Germany

Project Director

- Peter Shepherd, COUNTER

Usage Factor project

Acknowledgements

- Major funding from UKSG and RIN
- Stage 2 :additional funding from: ALPSP, American Chemical Society, STM, Nature Publishing Group, Springer
- Stage 3: additional funding from : American Association for the Advancement of Science, American Chemical Society, American Medical Association, American Institute of Physics, British Medical Association, COUNTER, Emerald, STM, Taylor & Francis, Wolters Kluwer

For further information

The full report on Stages 1 and 2 of the Usage Factor project, as well as the draft Code of Practice and information on the progress of Stage 3 is on the COUNTER website at:

http://www.projectcounter.org/usage_factor.html