
JUSP ebook discussion forum report

7th July 2016, Brettenham House, London

Contents

JUSP ebook discussion forum report	1
Present at meeting	2
Purpose of the meeting	2
JUSP context	3
Discussion	4
Library challenges	4
Publisher challenges.....	4
Common challenges.....	5
What currently works well (library perspective)?	5
Library wishlist	6
Other notes arising from discussion	6
Action plan and next steps.....	7
Appendix – Technical issues.....	8

Present at meeting

Alan Benton	Askews and Holts
Alison Brock	The Open University
Lorraine Estelle	Project COUNTER
Bettina Fischer de Vásquez	Springer
Sally Iannacci	Oxford University Press
Erica Lee	Aston University
Christian Linders	BIBSAM (online for part of meeting)
Anna Sansome	UCL
Maggie Sarjantson	University of Hull (also representing the Northern Collaboration)
Britt-Marie Wideberg	BIBSAM (online for part of meeting)

JUSP team

Jo Alcock	Evidence Base, Birmingham City University
Dave Chaplin	Jisc
Jo Lambert	Jisc
Paul Meehan	Jisc
Paul Needham	Cranfield University
Anna Vernon	Jisc
Laura Wong	Jisc

Apologies

Paul Harwood	EBSCO
Dahman Soltani	ProQuest

Purpose of the meeting

In order to progress the shared goal of developing, delivering and utilising consistent, reliable usage statistics for ebooks, Jisc is coordinating a representative group of publishers and librarians to discuss these issues. We hope that by surfacing and documenting current challenges and issues for all, we can work towards developing collective action plans to provide greater coherence and clarity in this area. The aim of this group is to provide a supportive environment to share issues from all perspectives and identify opportunities to support change.

The outcomes of the meeting are intended to:

1. Support development of COUNTER
2. Provide greater clarity around use and analysis of book reports
3. Explore opportunities for greater optimisation of services and support mechanisms

4. Develop a shared understanding of challenges from all perspectives

JUSP context

To provide some further context to the discussion and the motivations for convening the forum, the JUSP team shared their experiences with ebook usage data so far.

Jo Lambert started by talking about development of the JUSP ebook service which was released 29th February 2016. JUSP for ebooks currently includes data from 10 publishers supplying Book Report 2 and Book Report 3, however participating libraries want to see all their ebook usage data in JUSP.

Next, Jo Alcock gave an overview of ebook baseline survey conducted in May 2015 by the JUSP team. The report is available on the JUSP website <http://jusp.mimas.ac.uk/news/JUSP-ebooks-survey.docx>. Reasons for collecting ebook usage data included: assessing value for money, reporting to management, informing renewals, reporting to SCONUL, and understanding longitudinal trends.

Key issues for libraries were:

- » Not all publishers are COUNTER compliant
- » Time consuming process to collect the data
- » Inconsistency of data provision i.e. only offering BR1 or BR2, not both

Paul Needham then explained the JUSP process for collecting ebook usage data, and highlighted the technical issues the JUSP team have encountered.

- » No COUNTER Release 4 SUSHI services available for some publishers / aggregators
- » Reports that do not conform to Release 4 Code of Practice, and the SUSHI Implementation Profile
- » Reasons for non-compliance include: internal vendor IT, and lack of clarity in the SUSHI documentation
- » There are significant issues and minor issues (see appendix for details). Missing identifiers and multiple ISBNs were discussed in particular.

Discussion

Vendors and librarians split into two separate groups to discuss the current challenges within their organisation around creating, managing, gathering, analysing or reporting ebook usage data. The salient points from these discussions were then reported to the wider group.

Library challenges

- » Bringing together both COUNTER and non-COUNTER reports with information such as subjects, purchase models and reading lists is often very difficult, time consuming and inaccurate due to lack of key common identifiers and accurate entitlement lists. Librarians need this additional information in order to analyse and interpret the usage statistics meaningfully.
- » Difficult to track availability of titles and their usage. Comparing usage when access to the same title is available on different platforms. Identifying titles that have been dropped from packages or platforms, or where entitlements/business models change. Do we need to buy more user licences?
- » Difficult to bring together data from different sources due to lack of common identifiers in usage reports and title lists
- » Lack of accurate title-package information. What can our users access/use?
- » Lack of consistency over terms for purchase models
- » Want to be able to show how many times a title has been accessed but neither BR1 nor BR2 (nor a combination) are fit for this purpose. This makes what seems like a simple request much more complex
- » Ambiguity of the "section" definition, and inconsistency in its application
- » Benchmarking/usage profiling is difficult due to inconsistencies in reporting of BR1 and BR2. This makes comparing statistics challenging and means there are numerous caveats which need to be understood by anyone it is shared with

Publisher challenges

- » Business models versus COUNTER reports e.g. how to show usage for different types of purchase and different types of use? How to provide subject area reporting?
- » Publishers/vendors provide custom reports with collection or entitlement in order to provide libraries with the information they need. These custom reports may or may not use COUNTER statistics
- » Un-owned content (e.g. Patron Driven Acquisition). How is this identified in usage reports?

- » Multiple ISBNs for different incarnations of a book e.g. print book, ebook, online book. There is no way of reliably matching the “same” book on different platforms.
- » What is a book section? Different meaning for each vendor, and between different item types from the same vendor.
- » Libraries ask for the BR2 because they want to compare and report to SCONUL. However, this leads to further inconsistencies and inflated statistics (e.g. multiple section counts for a single download). Publisher and aggregator data are not comparable as they use different section requests.
- » Difficulties working with discovery platforms, sharing records and data and keeping up with content changes
- » Printed from online viewer – does this count?

Common challenges

It was clear that the concerns and challenges from both vendors and librarians were very similar.

- » Section requests inconsistent and BR2s are not comparable
- » What is a ‘section’? When should the BR1 or BR2 be used? Why not both? A single platform can support multiple types of access.
- » Bringing together usage data with business models, entitlements, subject area and print information due to lack of a reliable identifiers and knowledge bases
- » Tracking titles across different business models and platforms due to lack of a master identifier

What currently works well (library perspective)?

- » Using reports from one provider to view usage of different items from a single provider or compare usage over time periods (i.e. viewing reports in isolation)
- » High level aggregated usage trends are less affected by the variations in counting across providers (i.e. section definitions) and key trends over time can still be identified (e.g. annual increase in usage or peaks at certain times of year)
- » The Book report 3 show what content users are trying to access and demonstrate demand for additional titles or user licences and so good for informing purchasing decisions
- » Patron Driven Acquisition (PDA) is a purchasing model where access to a collection is made available to users but the library only purchases titles if a specified level of accesses is reached. It is a usage driven acquisition method and is popular as it demonstrates demand.

- » JUSP helpdesk improving accuracy for all by carrying out routine checks on the data and working with publishers to resolve issues.
- » SUSHI collection gives more time for analysis. Although currently there only a relatively small number of ebook vendors providing a compliant SUSHI service, this offers huge time saving benefits and is an integral part in the provision of ebook usage data.

Library wishlist

- » Greater transparency over 'section'
- » Consistency
- » Reference items to be listed separately
- » Data mining excluded
- » Clarity
- » Count 'accesses'

Other notes arising from discussion

- » COUNTER should ideally focus on what is being used, and not concern itself with entitlements
- » BUT there is a need for quality knowledge bases and assured sources, and for these to work with the reports
- » There are many issues associated with identifiers, and this seems to be the biggest barrier
- » Greater transparency, consistency and clarity is needed

Action plan and next steps

The group was supportive of actions currently being undertaken by the JUSP team:

- » To support JUSP participating publishers and aggregators planning to implement or test SUSHI services
- » To provide feedback on implementation challenges to COUNTER to support Release 5 and to USUS
- » To develop materials to assist JUSP libraries in understanding and using COUNTER book reports via JUSP
- » To conduct a usage statistics research project (a proposal was circulated ahead of the meeting)

Actions arising out of the discussion:

- » ACTION: vendors to share templates of their custom reports with JUSP and COUNTER
- » ACTION: libraries to share templates of their analysis of ebook data with JUSP and COUNTER
- » ACTION: MS/JUSP to continue to pursue opportunities for collaborative work with the Northern Collaboration
- » ACTION: details of subject identifiers used by publishers present to be shared and reviewed
- » ACTION: JUSP to involve platform providers, e.g. Highwire, SemanticO in future discussions around ebook usage
- » ACTION: JUSP to extend the planned research project to gather information about processes. There was a consensus that more detailed information is required around how vendors and libraries are utilising ebook data as part of existing processes and how it's being used to support decision-making
- » ACTION: Issues and questions raised around SUSHI and COUNTER to be logged by JUSP and used to help inform development of COUNTER Release 5
- » ACTION: Askews and Holts to send ebook account information to JUSP

Appendix – Technical issues

Significant issues (show stoppers)

- » XML does not match the SUSHI schema
- » ItemIdentifier labels
 - › ISBN must be Print_ISBN or Online_ISBN
 - › ISSN must be Print_ISSN or Online_ISSN
- » ItemIdentifier values
 - › ISBNs must be formatted with hyphens or spaces
- » Total for all titles
 - › totals must not be included
- » Begin and End dates not in YYYY-MM-DD format
- » Empty reports returned without any exception messages
- » Erroneous exception messages

Minor issues (that can be worked around)

- » MetricType
 - › Instances with a zero Count should be omitted
- » ReportItems
 - › with zero usage should be omitted
- » ItemIdentifiers
 - › With empty or n/a values should be omitted
- » We can ignore these but they make reports bigger
 - › Waste bandwidth
 - › Take longer to process

JUSP issues

- » We need to accurately identify items to:
 - › Provide quality assurance
 - › Match items across publishers and aggregators
 - › Facilitate interoperability with other services, e.g. KB+
- » But ItemIdentifiers in reports are only mandatory if available
- » Lack of identifiers doesn't break the COUNTER rules, but it does make life difficult for us!