

## Usage Statistics at Knowledge Unlatched

### Background

For digital products, publishers provide libraries with usage data in the form of COUNTER reports. In the evaluation process of research material held by a library, librarians also make use of other means of identifying topical areas and products of high or low usage as important to their collection development plans.

Moreover, there is an entire universe of tools and services that aim at supporting academic institutions in evaluating the impact of their research both within academia and beyond. For university administrators, library statistics and the insight derived from it have become important data points to measure the effectiveness of budget expenditures. Based on a set of statistics, they may allocate financial resources to libraries – or shift them elsewhere.

### Statistics and Open Access

Open Access publications pose a special challenge to publishers and librarians. While the basic service of COUNTER reports works well for users logged into the IP range of an academic institution, it is of significantly lower value for users of content that is freely accessible. And as approximately 75% of searches within library systems are performed from general search engines and not from Online Public Access Catalogues (OPAC), the likelihood of missing usage is very high.

Additionally, one could argue that measuring access to Open Access books and journals is somewhat contradictory to the idea of Open Access itself. It is, however, understandable that university administrators want to measure the effectiveness and efficiency of budgets not only for content bought in traditional business models, but also for Open Access.

### Measuring Usage at Knowledge Unlatched

Knowledge Unlatched, as a cooperative approach to freeing up research content through a global crowd-funding initiative, has been working on statistics from its Pilot phase in 2012. From its inception, libraries as well as publishers have paid attention to how the model works compared to their respective traditional models.

Based on data from the Pilot of 28 books spanning multiple sub-disciplines of the Humanities and Social Sciences, a team of researchers at Curtin University, Australia, under the supervision of Associate Professor Lucy Montgomery, has established a pragmatic and consistent set of statistics.

In contrast to traditional publishers, Knowledge Unlatched at this point has some specifics that impact usage data. Firstly, KU does not host electronic versions of books in its own environment, but on OAPEN and HathiTrust as well as on The Internet Archive, the British Library and elsewhere. In many cases books are also on publisher platforms (where available). However, for technical reasons, the statistics provided so far only count usage from OAPEN and thus reflect most likely only a fraction of the actual use.

Knowledge Unlatched provides, as specified for all content on OAPEN, books as one document for full-text downloads. In the representation of statistics, KU does not make any adjustments to mirror the industry standard of chapter downloads as some publishers do. To reflect chapter downloads and hence compare directly to publishers' statistics, the download figures would need to be multiplied by the number of chapters (ten to fifteen in most cases).

## **Library Statistics**

For librarians, full-text downloads are the “hard currency” of their spend on monographs, providing Return on Investment metrics. The KU library dashboard allows librarians to identify not only usage per title for their individual institution, but also the overall download performance of all libraries worldwide down to title level. If needed, librarians can order their individual statistics in Excel for further analysis, and they should keep in mind that right now, Knowledge Unlatched provides data only for titles hosted on the OAPEN platform. Downloads from HathiTrust and other hosts, as well as publishers’ websites are not included. Only usage by patrons logged into the library system are counted.

## **Publishers’ Statistics**

Publishers’ statistics include downloads per title per quarter in aggregate as well as on an institutional level. Publishers will be able to assess the performance of each of their own titles and compare performances. Publishers will only see usage data of their own titles, but this can be compared with anonymised and aggregated data from the full KU collection.

## **Next Steps**

From early 2017, Knowledge Unlatched will provide statistics to libraries as well as publishers on a regular basis. They will be available in PDF to all participating libraries through the KU portal.

## **Non-Institutional Usage**

As a large amount of usage potentially happens from users not linked to IP ranges of libraries, Knowledge Unlatched is working on making this usage visible on a geographical level. A participating institution will therefore also see non-registered usage from its metropolitan area. Hence, libraries can make a better approximation of which usage they assume to be attributable to their financial commitment.

## **Privacy of Data vs. Open Science**

Knowledge Unlatched respects the highest standards in privacy protection. KPMG, a leading global compliance firm, has reviewed our processes and certified that Knowledge Unlatched does not have control over individual users’ data in any part of its statistics process. Knowledge Unlatched can therefore only analyse and share data on a title, institutional, and/or publisher level.

In order to support Open Science and Open Data projects around the world, Knowledge Unlatched makes its data available to researchers for non-commercial work.

June 2016

Dr. Sven Fund  
Managing Director  
Knowledge Unlatched

[sven@knowledgeunlatched.org](mailto:sven@knowledgeunlatched.org)