

## Populating and Synchronizing Serials Solution Resource Manager 360 with SFX data: Experiences from the Field

Tony Harvell  
*University of California San Diego, tharvell@ucsd.edu*

Follow this and additional works at: <https://docs.lib.purdue.edu/charleston>

An indexed, print copy of the Proceedings is also available for purchase at:

<http://www.thepress.purdue.edu/series/charleston>.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: <http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences>.

---

Tony Harvell, "Populating and Synchronizing Serials Solution Resource Manager 360 with SFX data: Experiences from the Field" (2009). *Proceedings of the Charleston Library Conference*.  
<http://dx.doi.org/10.5703/1288284314802>

## POPULATING AND SYNCHRONIZING SERIALS SOLUTIONS RESOURCE MANAGER 360 WITH SFX DATA: EXPERIENCES FROM THE FIELD

*Tony Harvell - Head of Acquisitions, UCSD Libraries*

University of California Libraries has a long history of managing electronic resources collectively going back to the 1980's. UC Libraries represent ten campuses plus the California Digital Library (CDL). UC Libraries manage most electronic resources consortially through the California Digital Library, a "co-library" of the University of California based at the University of California's Office of the President. CDL Acquisitions are operated from UCSD Libraries. Though each campus has its own ILS, UC Libraries share a common union catalog (Melvyl) currently use Aleph software and an open URL Link Resolver (UC E-links) which operates using ExLibris' SFX. Collection development is coordinated system-wide for both print and electronic resources through all-campus groups.

UC had an unsuccessful joint- development project with ExLibris for a consortial version of Verde in 2007-2008, which was abandoned. In early 2010, CDL went into production with Serials Solutions Resource Manager 360, with much of the planning done in 2009. CDL Acquisitions also maintains the database custom dates for titles in Resource Manager. UC Libraries adopted SFX as their Open URL Link Resolver in 2003 with a shared instance of SFX across the ten campuses and 2 national labs managed by CDL. CDL Acquisitions centrally maintains the SFX instance for resources licensed for all or most UC campuses. Each campus library maintains a local instance for those things licensed locally. SFX data is used to populate Libraries "Request" – a VDX interlibrary loan tool implemented in 2002 and this component is also closely tied to UC's pilot of WorldCat Local for a future union catalog which began in 2009. Consequently UC has a significant commitment to continuing with SFX as a link resolver.

Initial data population of Resource Manager took six months undergo quality assurance. Even though we tried to use machine readable data whenever possible to batch load, we found that the inherent incompatibilities between disparate data sources made it very difficult.

We used exported data from four existing data sources:

- 1) SFX titles/thresholds were extracted and formatted in the Offline Date Status Editor format as proscribed by Serials Solutions for data loading. These were exported target by target from SFX and sent to Serials Solutions who had contracted to input the data for us.
- 2) Cost record data came from Order records in the III Millennium Acquisitions module, exported in delimited format. Because we used a customized approach to storing complex payment information for multiple campuses in the NOTES field, we had to input most of this data manually. We only input the current and previous fiscal years, with plans to go back retrospectively when time permits.
- 3) License data covering about 30 fields that had been analyzed over the last two years and input into Excel was cut and pasted into RM360 (no data loading program existed)
- 4) Contact data was loaded from a Microsoft Access file developed over several years and matched to database name and loaded manually.

As of November 1, 2009 the CDL implementation reported the following statistics.

Databases	486
Unique journal titles	75,017
Cost records (current FY only)	336
License records	119
Contacts	460

For the most part, UC Libraries do not normally license all titles in a given publisher package (i.e. the “Big Deal”) rather pick and choose titles based on value and cost per use. For the largest publisher packages; Elsevier, Wiley-Blackwell, Springer, Taylor and Francis, Sage – we have a formal evaluation process during each contract to determine the content each year. Consequently, we could rarely activate a full package in RM360 and do very little maintenance.

From the inception, we realized we would have difficulty reconciling data that originates from different data sources and we tried various ways to compensate for these differences. The major problems we identified included:

- 1) Serials Solutions does not track volume/issue enumeration – unlike SFX. Therefore, any data in Serials Solutions that would be automatically exported into SFX would lose that granularity.
- 2) The target names in SFX and Serials Solutions Knowledge Works database names do not always match. Though both producers aspire to package the data in the ways it is sold to customers, they don’t always match and certainly don’t have the same names. Through trial and error, we often found that we had chosen the wrong database to activate in Serials Solutions, and had to start all over again.
- 3) Some publisher packages/backfiles/archives are configured differently in the two KB’s. This is certainly true in cases where there are moving backfile/front file dates.
- 4) There are significant differences in how title changes are tracked in both KBs. Though both rely both on vendor data and cataloging data, how title changes are identified and captured will vary. This is particularly true of journals that are issued in parts or that have supplements.
- 5) SFX is updated monthly, Serials Solutions is updated more frequently but changes are difficult to identify unless one subscribes to their MARC service (UC does not).

Our initial analysis comparing the two KBs showed as much as 20% difference in title counts for some major publisher packages.

UC made several decisions based on available staffing and access considerations:

- 1) We decided to allow SFX to track aggregator content for link resolution – no coordination with RM360. We would not attempt to reconcile the two KBs because of the frequency of changes.
- 2) CDL had decided in 2008 that Serials Solutions would be the “database of record” for CDL licensed content for licensed publisher packages, its data would need to drive the SFX threshold, while Serials Solutions would be used track the entitlements. We decided to focus our efforts on publisher packages with the greatest subscription value. All KB maintenance focus would be on publisher packages with selectable content.
- 3) We would not use either KB to track monographic content. The KB for monographic titles has not been normalized by either provider, making it difficult to reconcile with

cataloging records. We did, however, activate databases with monographic content (e.g. CRCNetBase) for both SFX access and entitlement in Serials Solutions where we have access to the entire content. We normally do not activate targets for collections where we have selected titles (e.g. Ebrary, EBL, etc.) We do title level cataloging for most all monographs licensed by CDL or UCSD.

These were tough decisions because we knew that we would need to reconcile data across the system. California Digital Library technical staff worked with Serials Solutions to develop an SFX report to allow SFX users to make use of work done in RM360. This report is formatted specifically for the **SFX** data loader program to provide much greater efficiency and savings of time managing data. Report allows you to extract the Serials Solutions Journal ID and formats dates in the SFX parsed format Loader program. This would allow us to update data from Serials Solutions into SFX, but not the reverse.

Below are the fields of the SFX Report which is part of the RM360 suite of reports:

ISSN / ISBN  
ACTIVE (term required by SFX data loader)  
Parsed Date String (required by SFX data loader)  
Title (Serials Solutions authority title)  
Status (based on Resource Manager status)  
HoldingsID

This holdings ID is used as the unique match point to ensure integrity when updating from Serials Solutions to SFX. We did an initial load of 5000 titles to get the Holdings ID into SFX records in the "local field" in the Object record. This data is protected in KB updates done each month by SFX. Initial loads had about an 85% match rate. Non-matches have to be resolved manually. Often it is a discrepancy of ISSN (print versus e-issn), or of journals with supplements or multi-parts with separate ISSNs. SFX often handles these titles differently than does Serials Solutions. Fortunately SFX usually has both print and e-issns available which facilitates matching.

### **Process of managing the KB updates:**

Each month CDL Acquisitions staff run an SFX report in RM360 on selectable databases to identify changes that need to be made in SFX. Normally the list of databases is determined by the number and kind of changes that show up in the KB updates that come from SFX. Not all packages are done each month. We have found we need to do this particularly when the title content is being revised or titles are added or lost from a package.

Once the SFX report is extracted, we use the SFX Data loader Tool to prepare the update for the SFX instance. For those updates that are not made (usually because of inadequate matches by Holdings ID or ISSN), we need to review them to determine if ISSN, title, or holdings changes need to be reported to either Serials Solutions or SFX.

We established the following priorities for updating publisher packages:

Elsevier  
Wiley Blackwell  
Springer  
InformaWorld PLC  
Oxford

Cambridge  
Sage

Others are done as needed, depending on number of SFX KB Updates that may show up in each month. What we have found is that the uploading process does not work as well as we had hoped, and we have returned to doing manual comparisons of exported data between the two KBs to identify where adjustments can be made. The major issues are these: title/issn variations, missing ISSNs, and the inherent problem of not knowing when changes occur in the Serials Solutions Knowledge Works. It is possible for us to unknowingly update data in SFX that was correct with incorrect data from Serials Solutions. Consequently, we decided to manually review the data from SFX updates and compare it to Serials Solutions, making changes only when we have verified they are correct. We verify title changes, ISSNs and access and report discrepancies to either or both of the vendors. The SFX knowledgebase has much more flexibility for adding object portfolios instantly, whereas we have to wait for Serials Solutions to make the changes in their Knowledge Works and the change does not appear for at least 24 hours. Sometimes either of the knowledgebase producers will decline to make the change we recommend because of their particular policies. This has occasionally resulted in “irreconcilable differences” in data in the two systems.

There remain a number of outstanding issues we are looking to address. Programmers are working on protecting volume/issue enumeration in SFX if possible so that we don't lose that data with an upload from Serials Solutions. This is complicated because enumeration needs to be predicted from Serials Solutions data. We are also looking at improving programming to only update thresholds that have end dates added. This would eliminate any unnecessary changes.

In summary, we have had a mixed experience. When used with a limited number of resources, it works fairly well. It required working with both SFX and Serials Solutions to improve the data in their KBs. Hopefully all KBs can be improved upon through closer relationships between publishers and users of KBs – perhaps through groups such as The Knowledge Base and Related Tools (KBART) Working Group (NISO and UK Serials Group). Ultimately, vendors, publishers, and customers could benefit from having the knowledge bases maintained centrally, much like OCLC's WorldCat, so there is not so much maintenance. Vendors could focus their energies and sell products based on other features, and spend less time on KB maintenance. We understand that KBs are a valuable part of their products, but in essence consistent data, following established cataloging standards, and delivered in a standard format will benefit both libraries and vendors. As this part of the industry matures, we hope to see a greater move toward cooperation in the area of knowledgebase maintenance and design to allow more interoperability between link resolvers and data providers and electronic resources management systems.