

What do we need to build a successful knowledge base?

Maretha Allwright

SEALS Trust & Systems Manager

maretha.allwright@nmmu.ac.za

Dorette Snyman

Collection Developer: Commercial E-Resources, Unisa Library

snymad@unisa.ac.za

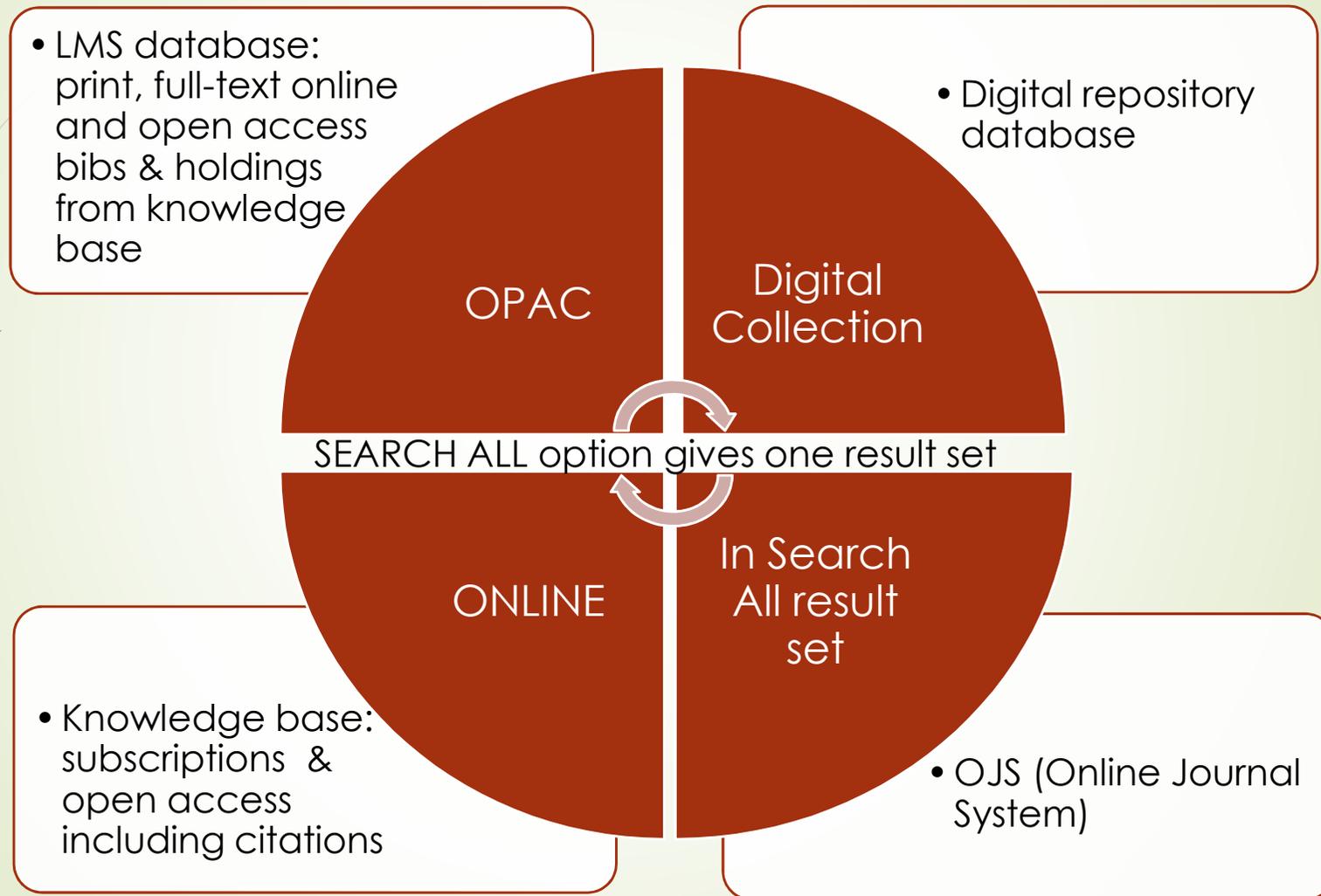
<http://orcid.org/0000-0001-9072-9820>



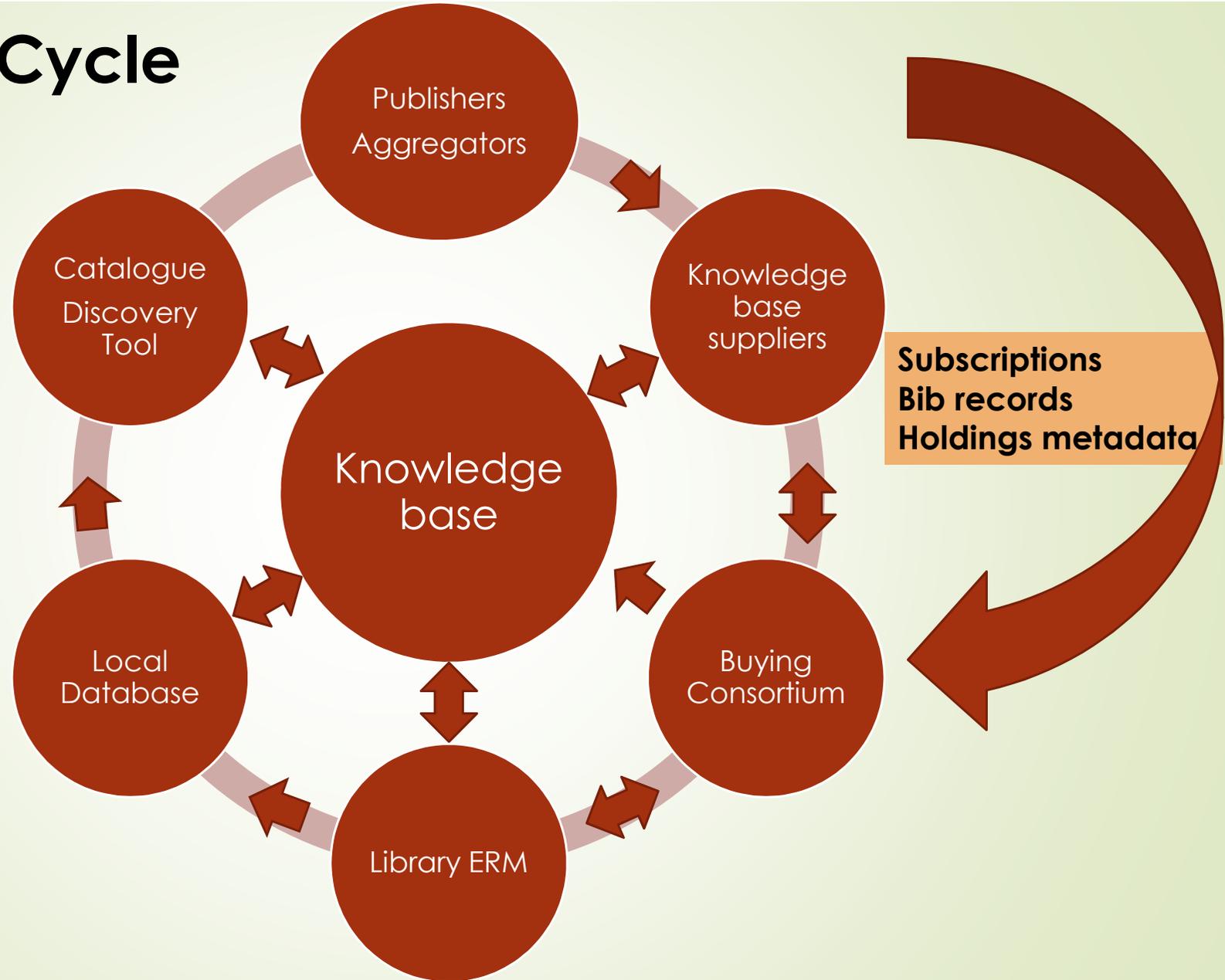
Overview

- ▶ Supply chain
- ▶ Metadata cycle
- ▶ What librarians need
- ▶ Why librarians need it
- ▶ What can be done for the future?

USERS - Output to Discovery Tool



Metadata Cycle





We rely on

- ▶ Correct and exact title lists of the consortium deals
 - ▶ not including site specific title subscriptions
- ▶ Correct and exact holding lists in KBART format of the subscribed titles
- ▶ Good RDA bibliographic records for subscribed titles
- ▶ Real-time and frequent updating of title and holdings information
 - ▶ By the supplier to the KB suppliers
 - ▶ When deals are renewed
 - ▶ When titles are added or removed during subscription period

KBART: Knowledge Bases and Related Tools Working Group



- ▶ KBART a joint initiative (with NISO) was launched in 2008 (<https://www.uksg.org/kbart>)
 - ▶ Phase I explored data problems within the OpenURL supply chain
 - ▶ They recommended best practices for formatting and distributing title lists (2010)
 - ▶ This enables content providers to greatly increase the accessibility of their products.
- ▶ KBART II is managed by NISO
 - ▶ Report on recommended practices in 2014 <http://www.niso.org/publications/rp/rp-9-2014/>
 - ▶ Section 3 of the report reports on “Metadata for Consortia”.
- ▶ The KBART support process includes a framework for enabling content providers to achieve compliance and publicly endorse their commitment to providing high quality metadata to the supply chain
- ▶ Promotes smoother interaction between members of the knowledge base supply chain: content providers, knowledgebase providers and libraries



KBART files

- ▶ Should be generated for all available collections sold by a content provider
- ▶ including Consortia deals if they differ from a standard collection
- ▶ They should exclude single title subscriptions by specific libraries
- ▶ Errors and inconsistencies in KB data cause services to break down
- ▶ Metadata should be provided in plain text encoded as UTF_8
- ▶ Metadata should be supplied in a tab separated metadata file with column headers
- ▶ Fields should appear in the order presented in KBART recommendation 6.1.1

KBART for Serials

publicaton_title	title_url
print_identifier	title_id
online_identifier	embargo_info
date_first_issue_online	coverage_depth
num_first_vol_online	notes
num_first_issue_online	publisher_name
date_first_issue_online	publication_type
num_first_vol_online	preceding_publication_title_id
num_first_issue_online	acces_type ← OA
date_last_issue_online	
num_last_vol_online	
num_last_issue_online	

Deep Dive into KBART - NISO <https://www.slideshare.net/NASIG/deep-dive-into-kbart>



Package/database name

- ▶ Package/database names must include the consortium name if the subscribed content differs from the general package of the supplier
 - ▶ SpringerLINK (SANLIC)
 - ▶ ScienceDirect Freedom Collection (SANLiC)
- ▶ Package name must be adjusted if new deals are signed under a new name OR new sub-packages are provided
 - ▶ New metadata must be supplied to the KB providers at the same time as the deals are negotiated and signed
- ▶ Package name must include the years of the subscription if specific periods are applicable

Packages in Knowledge Bases

PACKAGE NAME	PROVIDER
▶ ScienceDirect Freedom Collection (SANLiC)	Elsevier Science
▶ ScienceDirect Freedom Collection	Elsevier Science
▶ ScienceDirect Freedom Collection (Bibliosan)	Elsevier Science
▶ ScienceDirect Freedom Collection (NESLi2)	Elsevier Science
▶ ScienceDirect eBooks - Freedom Collection 2017	Elsevier Science
▶ ScienceDirect Freedom Collection (SwissUniCo)	Elsevier Science
▶ ScienceDirect Freedom Collection (Cristin)	Elsevier Science
▶ ScienceDirect Freedom Collection 2016 (DRAA)	Elsevier Science
▶ ScienceDirect Freedom Collection (NEICON)	Elsevier Science
▶ ScienceDirect Freedom Collection (BIBSAM)	Elsevier Science
▶ ScienceDirect Freedom Collection (JISC)	Elsevier Science
▶ ScienceDirect Freedom Collection (CCINSHAE)	Elsevier Science

Examples of title total differences in packages

➤ Freedom Collection (SANLiC)	2,195
➤ Freedom Collection (NESLI)	2,224
➤ Freedom Collection	2,361
➤ Freedom Collection (Cristin)	2,233
➤ Freedom Collection 2016 (DRAA)	2,262
➤ Freedom Collection (Bibliosan)	2,471
➤ Freedom Collection (Baden-Württemberg)	2,083
➤ Springerlink (SANLiC)	

1925 titles in KB (updated 12 May 2017 in KB)

2171 titles in Vendor List (Received 14/03/2017)

246 titles not in KB?

Business Source Complete

● Subscribed

Export ▾ Edit

SHOWING

Subscribed ✕

REFINE BY

Keyword ▾

Publication Type ▾

Serial (41,238)

Monograph (856)

Summary

Number of Titles 42,094
Subscribed 42,094
In Trial 0
Not 0

Subscribed
Content Type Aggregated Full Text

Source EBSCO
Package ID 1125
Package Type Available as a package
Last Updated 2017-05-12 17:13

Business Source Complete Magazines and Journals

6207 = Total number of journals & magazines indexed and abstracted (3457 are peer-reviewed)
3793 = Total number of journals & magazines in full text (1960 are peer-reviewed)

Last load had 6592 titles

Titles

- ▶ Variant titles caused by misspellings, title variations, or the use of previous or subsequent titles, lead to matching problems
 - ▶ Use full title as on print journal or homepage
 - ▶ Title changes should be listed separately with their set of coverage dates
 - ▶ Articles and subtitles should be included
 - ▶ Words or identifiers should not be added after titles
 - ▶ Series titles should not be added before proper titles
- ▶ Best practice
 - ▶ Use a form of the title that appears as a main title in a standard cataloging source
 - ▶ CONSER record set, the OCLC WorldCat database, or the ISSN Registry

Unique Identifiers: ISSN, EISSN, ISBN, EISBN

- ▶ The accurate use of the ISSN is critical in ensuring successful resolution of OpenURL-based links
- ▶ ISSNs should be correct and in the correct format, e.g. 0378-4363
 - ▶ (Ulrichs, ISSN portal, Library catalogs (LC, National Libraries))
- ▶ Do not re-use ISSNs for title changes, use print if only print available for earlier title
- ▶ Supply both the print and the online identifier if available and applicable to a specific publication
- ▶ Knowledge base developers should successfully manage multiple ISSN for a given title
 - ▶ Most titles that are available both in print and online have more than one valid ISSN assigned to them.



Dates, volumes and issues

- ▶ Use the ISO 8601 format for dates, e.g. 2013-05-14
- ▶ KBART does not support non-numeric dates like Spring, Summer etc.
- ▶ Start and closing dates must be applicable to the specific subscription (incl. consortium deals)
- ▶ Embargo dates must be applicable to a specific subscription (incl. consortium deals)
- ▶ Dates for titles that change or ceased publications must be correct
- ▶ Volume and issue entries must follow KBART recommendations and be applicable to the specific subscription (incl. consortium deals)

Start dates inconsistencies

Accident Analysis & Prevention

● Subscribed

Title Details

ISSN (print)	0001-4575
eISSN	1879-2057
Source	EBSCO
Publisher	Elsevier
Start Date	1969-07
End Date	To Present

Package Details

Package Name [ScienceDirect Freedom Collection \(SANLiC\)](#)

BUT SANLiC deal is only from 1995

Volume 28, Issue 2
pp. 139-279 (March 1996)

Volume 28, Issue 1
pp. 1-138 (January 1996)

Volume 27, Issue 6
pp. 741-857 (December 1995)

Volume 27, Issue 5
pp. 625-740 (October 1995)

Volume 27, Issue 4
pp. 425-623 (August 1995)

Volume 27, Issue 3
pp. 283-424 (June 1995)

Volume 27, Issue 2
pp. 143-282 (April 1995)

Volume 27, Issue 1
pp. 1-141 (February 1995)

Volume 26, Issue 6
pp. 689-820 (December 1994)

Volume 26, Issue 5
pp. 555-688 (October 1994)

Volume 26, Issue 4
pp. 411-554 (August 1994)

Volume 26, Issue 3
pp. 277-410 (June 1994)

Volume 26, Issue 2
pp. 135-275 (April 1994)

Volume 26, Issue 1
pp. 1-133 (February 1994)

Volume 25, Issue 6
pp. 647-791 (December 1993)

[Abstract](#) | [Purchase PDF](#)

Driver experience with antilock brake systems Original Research Article
Pages 807-811
Allan F. Williams, JoAnn K. Wells
[Abstract](#) | [Purchase PDF](#)

The seat belt syndrome in children Original Research Article
Pages 813-820
J.C. Lane
[Abstract](#) | [Purchase PDF](#)

Forthcoming papers
Pages I-II
[PDF \(154 K\)](#)

Volume 26, 1994 list of contents
Pages IV-X
[Purchase PDF](#)

Volume 26, 1994 author index
Pages XI-XII
[Purchase PDF](#)

Volume 26, 1994 keyword index
Pages XIII-XIV
[Purchase PDF](#)

When things go right (yes it can!)

- ▶ Journal *Sociologie du travail* cancelled its publication contract with Elsevier (ScienceDirect Freedom Collection) at the end of 2016
- ▶ Title automatically corrected in ScienceDirect Freedom Collection (SANLiC) with correct closing dates

Title Name				
Sociologie du Travail				
Package Name	Content Type	Coverage	Available Titles	Current Holdings
ScienceDirect Freedom Collection (SANLiC)	E-Journal	01/03/1999 - 12/31/2016	2,195	✓

- ▶ Title available for selection in the EBSCO EDS knowledgebase

Title Name				
Sociologie du Travail				
Package Name	Content Type	Coverage	Available Titles	Current Holdings
Revues.org	Aggregated Full Text	01/01/2017 - Present	450	✓

Entitlement list not in KBART format:

Inside OUP Unisa admin portal

↕ Name	↕ Start Date	↕ Expiry Date
African Affairs Jan 2004 to Dec 2009	1 Jan 1900	
African Affairs Jan 1999 to Dec 2000	1 Jan 1900	
American Literary History Jan 2004 to Dec 2007	1 Jan 1900	
American Literary History Jan 1999 to Dec 1999	1 Jan 1900	
Applied Linguistics Jan 2004 to Dec 2009	1 Jan 1900	
Applied Linguistics Jan 1999 to Dec 2000	1 Jan 1900	
Biometrika Jan 1999 to Dec 2000	1 Jan 1900	

OUP Librarians page

2017 KBART journals collection files

- [2017 Current collection](#)
- [2017 Archive collection](#)
- [2017 Current collection with Open Access](#)
- [2017 Economics and Finance](#)
- [2017 Humanities](#)
- [2017 Humanities Archive](#)
- [2017 HSS](#)
- [2017 Law](#)
- [2017 Law Archive](#)
- [2017 Life Sciences](#)
- [2017 Mathematics and Physical Sciences](#)
- [2017 Medicine](#)
- [2017 Medicine Archive](#)
- [2017 Open Access Collection](#)
- [2017 Policy](#)

Publishers providing correct holdings in KBART format

ScienceDirect

Electronic Holdings Reports

This is a new version of the Electronic Holdings Reports.
Click a link in the Available Reports area to download the Electronic H

Available Reports

These reports list your holdings information as of **17 May 2017**:

Account: SA Consortium - University of South Africa (UNISA) (COO

Select report format:

· **Comma Separated Value (CSV)**

[Download](#)

· **Microsoft Excel**

[Download](#)

· **XML** (usable as input format in e.g. OPAC systems)

[Download](#)

Institution

University of South Africa UNISA Library Serials Division |
2000292063

My SpringerLink access lists

My organisation's title list: Titles which your organisation has access to because you have a contract with us.

My consortia title list: Titles which your organisation has access to because you are a member of a consortium.

2000292063 | University of South Africa UNISA Library Serials Div ▾

My organisation's title list
2000292063 | University of South Africa UNISA Library Serials Division

My consortia title list
3000137117 | SANLIC

[Download Book list](#) [Download Journal list](#)

Unfortunately
SANLIC list is
empty



Why is a correct Kb so important?

- ▶ New Kb ingests the classic catalogue, IR, local and any other content collections libraries want to offer
- ▶ Kb is a core library infrastructure supporting most of the visible services delivered
- ▶ Libraries spend millions of Rand on subscribed online resources and should supply correct access to their users
- ▶ Users get despondent and loose faith in the system if they are offered dead links instead of the full-text articles or e-books
- ▶ Library staff cannot sift through thousands of titles to try and determine which titles or holdings are relevant to their subscriptions
- ▶ It is complex and time consuming to work with sub-standard metadata records and incorrect holdings data
- ▶ Accuracy and completeness of a library's kb fundamental requirement for quality services rendered
- ▶ Kb is central to build new generation services such as web-scale discovery and library services platforms (patron facing products)

Successful knowledgebase = manual labour?



All labor that uplifts humanity has dignity and importance and should be undertaken with painstaking excellence.

— *Martin Luther King* —

AZ QUOTES

Do you want robots to take over your job
(especially the manual labour part)?



What libraries want to see:

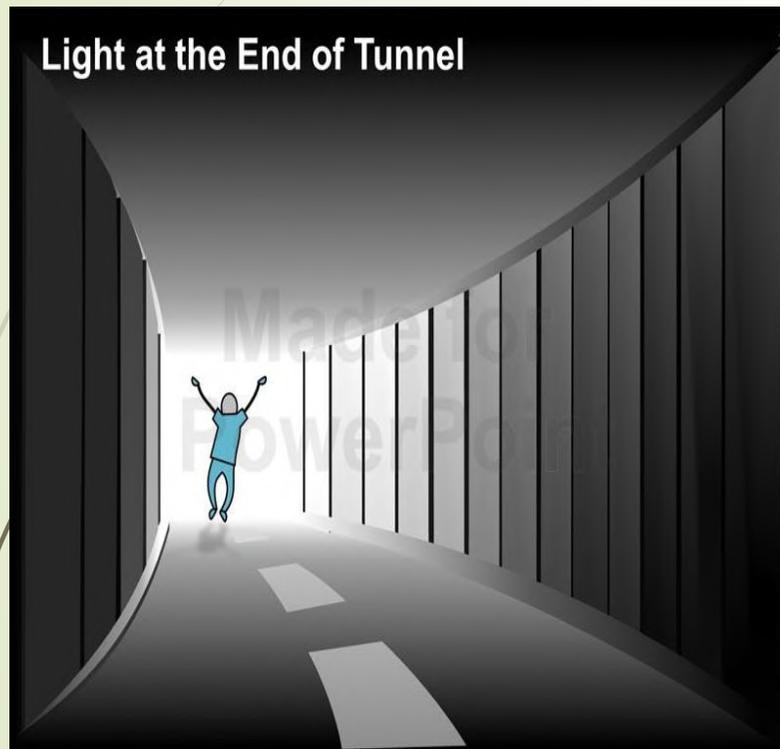
- ▶ Interoperability between systems
 - ▶ Leverage the technology imbedded in the information supply chain from content creators to Kb providers, through to discovery service to public availability
 - ▶ Support the increasing importance of web-scale discovery service as single access point of all resources including OA content
- ▶ Support efficient workflows in the metadata cycle
 - ▶ Reduce the number of investigative work required – checking, corrections, validation
 - ▶ Create collections in all Kb's equal to the collections offered (KBART II requirement)
 - ▶ Reduce troubleshooting of e-resources in support services



What libraries want to see:

- ▶ Supporting digital asset management and inventory control
 - ▶ Accurate and standardized holdings lists through the supply chain
 - ▶ Internal auditing of assets includes digital holdings as evidence of expenditure
 - ▶ Accurate at the time of offer to the consortium even for customized collections
 - ▶ After the “Big Deal” a record of retention of permanent content
- ▶ Support decision making – metrics & analytics
 - ▶ If users can't find it, they won't be using it – downward spiral of usage and funding
 - ▶ Quality knowledgebase can be used to support other services requiring holdings lists such as accreditation reports
 - ▶ Support collection management with derivate tools such as overlap analyses
 - ▶ Provide open access metadata (KBART II) will enable libraries to enlarge their content offering to end-users

A successful knowledgebase - is there light at the end of the tunnel?



<http://www.canstockphoto.com/illustration/tunnel.html>

OR





We hope that we can work together to make this work for South Africa!

THANK YOU