

Library Administration of Critical Research Infrastructure: Integrating a DSpace Repository & Research Information Management System Open Repositories, Denver, Colorado June 7, 2022

Megan Macken, Digital Resources & Discovery Services Clarke lakovakis, Scholarly Services & Research Engagement Oklahoma State University Libraries

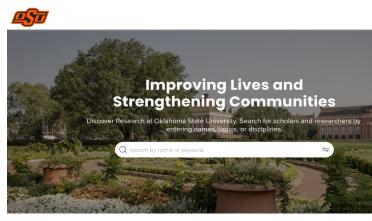
Download slides: <u>osf.io/9xpaq</u>

Research Information Management Systems

- A Research Information Management System (RIM System, aka CRIS or FAR) is a software platform for the collection, storage, and linking together of metadata related to faculty service, grants, teaching, and scholarly activities.
- A single point of access for data typically spread across multiple internal and external platforms
- Profiles are private by default but can be toggled to public and displayed on a public profile website
- Helps researchers, funders, and members of the public identify expertise within the university, and enable faculty members to communicate their research focus and activities more effectively
- Saves time and increases the accuracy of university reporting

RIM System Implementation at OSU

- Symplectic Elements implementation led by OSU Library from 2019 in partnership with IT and other university staff. We refer to Elements locally as **Experts Directory.**
- Publications metadata harvested from external platforms such as Scopus, Dimensions, Crossref, etc.
 - Faculty "claim" work based on author identifier or name searches
- Four other data feeds come from internal university sources:
 - Personnel information (Human Resources)
 - Grants (Office of Grants & Contracts Financial Administration)
 - Courses taught (Registrar)
 - University scholarly & creative works (SHAREOK)



https://experts.okstate.edu/



DSpace at OSU: SHAREOK

- Three universities/DSpace communities in the SHAREOK consortium
- Consortium impacts updates, system modifications, metadata
- DSpace 6; DSpace 7 anticipated in 2023
- Mediated deposits
- OSU Community Collections:
 - faculty papers, patents, conference proceedings, theses and dissertations, Ag Extension publications, etc

ing Okianoma scholarship	, research and institutional memory
REOK Home	
٩	SHAREOK Repository
HAREOK	SHAREOK is the joint institutional repository for the University of Oklahoma Libraries (OU), Oklahoma State University Libraries (OSU) and the University of Central Oklahoma Max Chambers Library (UCO). It serves as the home for the intellectual output of those institutions, such as: digital theses and dissertations, faculty publications, open access publications, open educational resources, institution-specific content and much more.
e Date	If you would like to place your works in the institutional repository, please contact your library: OU, OSU, or UCO. Digital assets place in the repository are available to a global audience, and search engine optimization techniques are used to increase their visibility to researchers and thus their impact on the global community.
15	If you are an OU student submitting a master's thesis or doctoral dissertation to the Graduate College via SHAREOK, please follow these instructions. For policies on embargoing your thesis or dissertation, please refer to the Graduate College Thesis/Dissertation Instruction Packetat. OSU students should follow the Thesis and Dissertation Guidelines posted in Carwas or email the Graduate College for more information.
INT	Communities in SHAREOK
	Select a community to browse its collections.
	Oklahoma State University [33853]
I Sister Resources (311)	The University of Oklahoma [44797]

https://shareok.org

By Serie

Herland

Oklahoma Cooperative Extension Service

- "Ag Extension": University/government effort to "extend" the university via science-based educational programs
- Topics from agriculture to youth development
- Numbered fact sheet series, agricultural experiment station research reports, newsletters
- As factsheets are updated and republished, authors change
- Published online



Prenaredness

Integration of Elements and DSpace

- The Scholarly Services position responsibilities include providing the OSU community with outreach and education on open access and the advantages of utilizing SHAREOK to archive and provide access to their work
- In 2021, the library initiated a project to integrate Elements and DSpace. We worked with Symplectic to guide the integration process and help construct the XML schema and crosswalk map files
- This integration works both ways:
 - Deposit: Allows researchers to deposit their work (articles, slides, posters, data, etc.) from Experts Directory to SHAREOK
 - Harvest: Imports metadata records already existing in SHAREOK, and allows researchers' to "claim" those records in their Experts Directory profiles.

Integration of Elements and DSpace

- The integration of these systems reinforces Elements as a consistent, centralized, curated database of scholarly and creative works across the university
- Implementation took roughly 6 months, and the system went live in January 2022.
 - SHAREOK deposit support page



Mapping Metadata: Crosswalk

E 4. Field mappings

DSpace Records +

5	~ 🖶 🕈 100% ▼ £ 5	% .0 .00 123×	Arial 👻	10 • B		* 1+ * P * GD		Σ -									
	• <i>f</i> X																
	A	В	с	D	E	F	G	н	- E	. U	К	L	м	N	0	P	
			Default field usage	Direction	Default field mapping (repository field)	Elmts. Pub'n type / Repo. Item type	Elements	Repository	Elements	Repository	Elements	Repository	Elements	Repository	Elements	Repository	
			(typically the most	<-> 2-way	This is the field mapping that will be used for	Display names ->		Article	Book	Book	Book chapter	Book chapter	Book review	Book review	Composition		C
-			common display		all item types, unless a different mapping is		journal-article	David Falls	book	Deces Cold	chapter	Dura Cald	c-book	D	composition	Den Calif	CO
	Underlying field in Elements	data type	name for the underying field)		explicitly specified for a particular type in the n columns to the right.			Repo. field k for type 'journal-a	Elmts. field for type 'book'	Repo. field for type 'book'	Elmts. field for type 'chapter'	Repo. field for type 'chapter'	Elmts. field for type 'c-book'	Repo. field for type 'c-book'	Elmts. field for type 'composition	Repo. field for type 'compos	iti ty
1	title	text	Title	(-) *	dc.title -	29 :: 4	[as default]		[as default]		 Chapter title 	[use default ma 🖛	Book review tit	le (use default ma =	[as default]		- [a
	abstract	text	Abstract	<-> *	[complex mapping - describe in a comment -		[as default]		[as default]		- [as default]		[as default]	-	[as default]		- [8
	authors	person-list	Authors	<-> *	[complex mapping - describe in a comment -	127.2.1.1.1.1.1	[as default]		[as default]		- [as default]	-	Book authors	(no mapping - e -			- [8
	author-url	Inu	Author URL	> *	dc.relation.uri -	The second se	[as default]		[as default]		 [as default] 		[as default]		[as default]		- [8
	editors	person-list	Editors	<-> *	dc.contributor.editor -		[as default]		[as default]		 [as default] 		[as default]		-		- 1
	series edition	text	Series	<-> * X *	dc.relation.ispartofseries *	5::1	the second second second		[as default]		 [as default] 		[as default]				1
	volume	text	Edition	x -	[no default mapping]		[as default] [as default]		[as default] [as default]		 [as default] 		[as default]				- 14
	pagination	pagination	Pagination	x -	[no default mapping] - [no default mapping] -	10 :: 1	and the second se		[as default]		 [as default] [as default] 		[as default] [as default]				- 4
- 15	publisher	text	Publisher		dc.publisher *	14 :: 1			[as default]		 [as default] 		[as default]				- 4
	publisher-url	url	Publisher URL	> *	dc.relation.uri -		[as default]		[as default]		- [as default]		[as default]		[as default]		- 6
	place-of-publication	text	Place of publication	(-) *	dc.coverage.spatial -	8::1	les servery		[as default]		- [as default]		[as default]		(as contrary		- 14
	publication-date	date	Publication date	(-) *	dc.date.issued	a second s	[as default]		[as default]		- [as default]		[as default]		[as default]		- 1
	isbn-10	isbn-10	ISBN-10		[no default mapping]	8::1			[as default]		- [as default]		[as default]	-			- 1
1	isbn-13	isbn-13	ISBN-13	х -	[no default mapping]	8 :: 1			[as default]		- [as default]		[as default]		5		- [1
	doi	doi	DOI	(-) *	dc.identifier.doi -	20 :: 1	[as default]		[as default]		- [as default]		[as default]				- [
	medium	text	Medium	х -	[no default mapping] -	20 :: 2	[as default]		[as default]		- [as default]		[as default]	,	[as default]		- [1
	publication-status	choice	Status	х -	[no default mapping] -	20 :: 2	[as default]		[as default]		+ [as default]		[as default]	-	[as default]		- [1
,	keywords	keyword-list	Keywords	(-) *	dc.subject.keywords -	29 :: 2	[as default]		[as default]		 [as default] 		[as default]	-	[as default]		- [4
	notes	text	Notes	x -	[no default mapping]		[as default]		[as default]		 [as default] 		[as default]		[as default]		- [8
	number	text	Number	<-> *	[complex mapping - describe in a commen -		and the second se	[no mapping - e -			 Chapter number 						-
	parent-title	text	Parent title	(-) *	dc.relation.ispartof	8::5					* Book title	dc.relation.ispa -	Book title	dc.relation.refer *			1
	name-of-conference	text	Name of conference		[no default mapping]	4::3	100000				*						- [2
	location start-date	text date	Location Start date	x - x -	[no default mapping]	9::7	Country	[no mapping - e *	-					-	-		- 0
	start-date finish-date	date	End date	x -	[no default mapping] -	8::4											- 0
	iournal	text	Journal	(-) *	dc.relation.ispartof -		[as default]	[use default ma -					[as default]				- P
- 1	issue	text	Issue	x -	[no default mapping]		[as default]	fore result that			-		fas colauri				- 1
	issn	issn	ISSN	x -	[no default mapping]	100000	[as default]				-						- 1
	pli	text	PII	x -	[no default mapping]		[as default]				-						-
-12	language	text	Language	(-> *	dc.language -		[as default]				-			,			
	patent-number	text	Patent number		dc.identifier.patentID -	1:1					-	-		-			-
- 17	associated-authors	person-list	Associated authors		dc.contributor -	1:1					-						
	filed-date	date	Filed date	4.4	dc.date.filed	22	Submission dat	12							-		-

2. Repository Field List ~ 3. Types mappings ~

🛄 🔒 Elements Field Grid (Processed) - 🔒 Repository Field List (Processed) - 🔒 Xwalk XML -Testing ~

4 >

Mapping Metadata: Types

A	В	С	D
Elements Publication Type (identifier)	Elements Publication Type (display name)	Repository Item Type (display name)	Repository Item Type (identifier, if different)
journal-article	Journal article	Article	
book	Book	Book	
chapter	Book chapter	Book chapter	
c-book	Book review	Book review	
composition	Composition		
conference	Conference proceeding	Conference proceedings	
dataset	Dataset	Dataset	
design	Design		
exhibition	Exhibition		
report	Report	Bulletin;Research Report; Technical Re	eport; Executive Summary
figure	Figure		
internet-publication	Internet publication	Blog	
c-newsletter	Newsletter	Newsletter	
artefact	Object		
patent	Patent	Patent	
c-internet-publication	Pedagogical material	Instruction	
performance	Performance		
c-media	Popular press		
poster	Poster	Poster	
preprint	Preprint	Preprint	
presentation	Presentation	Presentation	
scholarly-edition	Scholarly edition		
software	Software / Code		
thesis-dissertation	Thesis / Dissertation	Thesis;Dissertation	
media	Video/audio/media		
c-working-paper	Working paper	Working Paper	
other	Other		
fileset	Fileset		
c-report	Extension Publication		

- Team of four
- 20 +/- types
- dc.type.genre
- "Less is more" approach helps produce a more balanced workflow
- Omitted from auto deposit:
 - Video
 - Scholarly Editions
 - Some dissertations

osf.io/9xpaq

Mapping Metadata: Fields

- Authors
 - Dc.contributor
 - Dc.contributor.inventor
 - Dc.contributor.assignee
 - Reinman, Suzanne and Ahrberg, Janet (2020)
 "Issued Patents in a University's Institutional Repository," Journal of the Patent and Trademark Resource Center Association: Vol. 30, Article 5.
 <u>https://tigerprints.clemson.edu/jptrca/vol30/iss1/5</u>

OKSTATE Xwalk planner - Dspace Repository Q 0 File Edit View Insert Format Data Tools Extensions Help Last edit was 2 minutes ago 5000 100% + £ % .0 .00 123 + Arial 10 BISA & E = · T · - fx J17 D **Default field** Direction Default field mapping (repository field) 2 usage This is the field mapping that will be used for (typically the most <-> 2-way 3 common display <-- harvest all item types, unless a different mapping is **Underlying field** name for the --> deposit explicitly specified for a particular type in the N data type in Elements underying field) X no mappin columns to the right title text Title dc title 6-3 abstract Abstract text (-) [complex mapping - describe in a commen authors person-list Authors <-> [complex mapping - describe in a commen author-ur url Author URL dc.relation.uri --> editors person-list Editors dc.contributor.editor (-> 10 series text Series dc.relation.ispartofseries 11 edition text Edition [no default mapping] 12 volume text Volume Ino default mapping 13 pagination pagination Pagination Ino default mapping 14 publisher Publisher dc.publisher text publisher-url 15 url Publisher URL dc relation uni --> 16 place-of-publication text Place of publication (-> dc.coverage.spatial 17 publication-date date **Publication date** 6-3 dc.date.issued 18 isbn-10 isbn-10 ISBN-10 [no default mapping] 19 isbn-13 isbn-13 ISBN-13 Ino default mappingi 20 DOI do doi <-> dc.identifier.doi 21 medium text Medium [no default mapping] 22 publication-status choice Status [no default mapping 23 keywords keyword-list Keywords dc.subject.keywords 24 Notes Ino default mapping] notes text 25 number text Number [complex mapping - describe in a comme 26 parent-title text Parent title <-> dc.relation.ispartot name-of-conference text Name of conference [no default mapping] Ino default mapping 28 location text Location date Ino default mapping 29 start-date Start date 30 finish-date date End date Ino default mapping 31 iournal text Journal dc.relation.ispartof <-> 32 text [no default mapping] issue Issue 33 issn ISSN [no default mapping] issn 34 text [no default mapping] pi language text Language dc.language patent-number text Patent number dc.identifier.patentID person-list associated-authors Associated authors x dc.contributor filed-date date · dc.date.filed Filed date (->

3. Types mappings

E 4. Field mappings

DSpace Records

2. Repository Field List ~

+ =



Mapping Metadata: Fields

- Subjects
 - Keyword duplication
 - MESH
- Rights statements & licenses
 - → Dc.rights
 - Dc.rights.license = Author license
 - Publisher license

OKSTATE Xwalk planner - Dspace Repository * @ @ File Edit View Insert Format Data Tools Extensions Help Last edit was 2 minutes ago

17	• <i>f</i> X							
	A	8	С	D		E		
1			Default field usage (typically the most	Direction <-> 2-way < harvest > deposit X no mappir				
3	Underlying field in Elements	data type	common display name for the underying field)					
5	title	text	Title	<->		dc.title		
6	abstract	text	Abstract	(-)	*	[complex mapping - describe in a commen		
7	authors	person-list	Authors	(->	*	[complex mapping - describe in a commen		
8	author-url	url	Author URL	>	*	dc.relation.uri		
9	editors	person-list	Editors	(->		dc.contributor.editor		
0	series	text	Series	(->		dc.relation.ispartofseries		
1	edition	text	Edition	x	*	[no default mapping]		
2	volume	text	Volume	x	-	[no default mapping]		
3	pagination	pagination	Pagination	x	-	[no default mapping]		
4	publisher	text	Publisher	(->		dc.publisher		
5	publisher-url	url	Publisher URL	>	*	dc.relation.uri		
6	place-of-publication	text	Place of publication	(-)		dc.coverage.spatial		
7	publication-date	date	Publication date	<->		dc.date.issued		
8	isbn-10	isbn-10	ISBN-10	x	-	[no default mapping]		
9	isbn-13	isbn-13	ISBN-13	x		[no default mapping]		
0	doi	doi	DOI	<->		dc.identifier.doi		
1	medium	text	Medium	x		[no default mapping]		
2	publication-status	choice	Status	x	+	[no default mapping]		
3	keywords	keyword-list	Keywords	(->	*	dc.subject.keywords		
4	notes	text	Notes	x		[no default mapping]		
5	number	text	Number	(->	*	[complex mapping - describe in a commen		
6	parent-title	text	Parent title	<->	-	dc.relation.ispartof		
7	name-of-conference	text	Name of conference	x	*	[no default mapping]		
8	location	text	Location	x	-	[no default mapping]		
9	start-date	date	Start date	x		[no default mapping]		
0	finish-date	date	End date	x		[no default mapping]		
1	journal	text	Journal	(->		dc.relation.ispartof		
2	issue	text	Issue	x	*	[no default mapping]		
3	issn	issn	ISSN	x	*	[no default mapping]		
4	pii	text	PII	x	-	[no default mapping]		
5	language	text	Language	<->	*	dc.language		
6	patent-number	text	Patent number	<->		dc.identifier.patentID		
7	associated-authors	person-list	Associated authors	x	*	dc.contributor		
88	filed-date	date	Filed date	(-)	-	dc.date.filed		

1 2. Repository Field List ~ 3. Types mappings ~ 🗳 4. Field mappings ~

Workflow for Processing Deposits

=	<u>es</u> t		
2 2	C RETURN TO THE PREVIOUS PAGE Deposit scholarly & creative work: Previous Page Management Systems	rivacy and Resear	ch Information
מ	You are about to deposit this presentation to SHAREOK		>
	Deposit advice	1. Prepare deposit (step 1 c	f 3)
	Institutional advice Use this function to deposit your scholarship directly to <u>SHAREOK</u> , a repository for scholarly work authored by OSU community members. SHAREOK is managed and maintained by the OSU Libraries. For more information on this process, including the use of SHERPA ROMEO tab to review publisher policies, see this <u>Experts</u> <u>Directory support page</u> . Maximum file size is 1GB. For larger files, please contact experts directory@okstate.edu.	Upload a file Choose a file from your local mach Choose file: Browse N File version: [None]	Add OA location ine: io file selected.
			Leave without depositing

- Testing with Symplectic: iterative
 - Data structure & workflow changes
- User workflow:
 - File upload
 - Share an open access link



Workflow for Processing Deposits

	<u>eş</u> j			
C	RETURN TO THE PREVIOUS PAGE Deposit scholarly & creative work: P Management Systems	rivacy and	Research	n Information
	You are about to deposit this presentation to SHAREOK			>
	Deposit advice	1. Prepare dep	osit (step 1 of 3	3)
	Institutional advice	Upload	a file	Add OA location
	Use this function to deposit your scholarship directly to <u>SHAREOK</u> , a repository for scholarly work authored by OSU community members. SHAREOK is managed and maintained by the OSU	please enter the lo		e elsewhere in Open Access form,
	Libraries. For more information on this process, including the use of SHERPA RoMEO tab to review publisher policies, see <u>this Experts</u> <u>Directory support page</u> .	* OA location: File version:	~	
	Maximum file size is 1GB. For larger files, please contact experts.directory@okstate.edu.			Save location
				Leave without depositing

- Testing with Symplectic: iterative \bullet
 - Data structure & workflow Ο changes
- User workflow:
 - File upload Ο
 - Share an open access link Ο

Workflow for Processing Deposits

- Library workflow:
 - Increased production, add processes, workflow changes
 - check for an existing publication
 - move from the Development collection to Faculty Papers or another collection [NEW]
 - edit metadata to conform to local standards
 - \circ verify open access
 - re-upload bitstream with standardized filename [NEW]



Ag Extension Publication Deposits

- Concurrent metadata remediation process and transition to Open Journal Systems (OJS) for publication review
- This provided a holistic look at the extensive collection as we encountered issues during the harvest and deposit process:
 - Curated collection now with faculty-contributed metadata.
 - Who can edit data?
 - Locking the Extension Publication type in Elements
 - Working with Ag Communications Staff
 - Priority processing in DSpace for faculty papers



Versioning

- A single item might have multiple versions of the same object that must be identified
- Journal article versions
 - Users can specify different versions of a journal article (submitted, accepted, or published), which is added to the DSpace metadata
 - However, if the bitstream is updated (e.g. to change the file name), the version information must be re-entered because it pertains only to the version/bitstream



Versioning

- Ag Extension fact sheet versions
 - Updated with new research or as faculty are hired/retire
 - In DSpace we make new versions most visible, and make older versions available as distinct records
 - Even though each item has a distinct UUID, Elements only surfaces the most recent version and we must send older versions in a separate process.
- We opted to import all fact sheet versions because senior faculty may want to have that publication in their profiles, even if the information in the fact sheet has been superseded
- These are fundamental questions related to the purpose of the system: is it a tool for public profiles, annual reporting, or both?
 - Since it is currently both for us, we continue to deduplicate and support claiming older versions; however, this creates more work for the library related to file and version management

Version History

Version	Item	Date
4	11244/49846.4*	2022-04-28T20:26:40Z
3	11244/49846.3	2022-04-28T20:26:12Z
2	11244/49846.2	2022-04-28T14:38:12Z
1	11244/49846	2017-03-29T16:29:20Z



Deduplication

- Elements regularly monitors DSpace for new items; if it finds a matching item, the repository metadata is added to the publication object as a new record, otherwise Elements creates a new publication.
- In most cases, the automatic deduplication process in Elements is helpful, but occasionally it introduces errors that require us to split records;
 - Conference presentation merged with a journal article
 - An older fact sheet edition merged with a new fact sheet
- Deduplication can be time-intensive and generally requires manual verification



Results

- Increased usage of SHAREOK despite relatively soft marketing launch
- Cleaner metadata & less remediation for SHAREOK records
- Examining the mappings so closely led us to review the consistency and completeness of existing metadata in SHAREOK, as well as our data structure in Elements, leading to improvements in both systems



Next steps

- The technological barrier is lowered, but there is still work to do on policy side to incentivize and/or require deposit of manuscripts
- User education: author rights & versions
- Managing OA workflow between departments
- Upgrade to DSpace 7 projected for 2023
- Analysis & visualization of SHAREOK objects



Questions

Megan Macken

Oklahoma State University Libraries Assistant Department Head, Digital Resources & Discovery Services megan.macken@okstate.edu

Clarke lakovakis

Oklahoma State University Libraries Director, Scholarly Services & Research Engagement clarke.iakovakis@okstate.edu

View OKState public profiles at <u>https://experts.okstate.edu/</u>

Download slides: osf.io/9xpag

osf.io/9xpaq