### The research publications workflow at Deakin University Or, our transition from a manual to an automated workflow Prepared for the CAUL Repositories Community Day June 2017

### About Deakin

- Young institute around 50 years old

- @1500 researchers working across 4 campuses, 4 faculties, 4 research institutes and 13 strategic research centres
- Expanding our research capabilities and reputation
- One of Australia's fastest growing research universities research funding increase from A\$4.5 million in 1997 to A\$47.2 million in 2015
- 2015 ERA rankings: 89% of Deakin research rated at or above world standard
- 2015 Times Higher Education 100 Under 50 ranked Deakin University 45th in the World

### **Deakin Research Online**

- Created in 2008 / uses Fez/Fedora software
- Mixed repository of open access and dark archive material
- Broad collection policy based on the philosophy of "we'll take anything created by a Deakin researcher"
- Publications source of truth > publication data fed from the repository to Research Master then onto Staff profiles
- Strong relationship between Faculty, the Library, and the Research Office

Our publications workflow from 2008-2014 was based on manual processes involving record by record review of publication metadata and outputs



to Staff profiles

Deakin Research Online (Fez/Fedora)

### Introduction of Elements, late 2014

- The Library is the business owner of Elements
- Elements creates profiles for researchers, then harvests publication metadata from sources such as Scopus, WoS, etc.
- Metadata is fed through to DRO, then onto RM and Staff profiles
- We adapted the HERDC module to control workflows based on the old model of manual review
- Searches were turned on in late 2014, and Elements began harvesting data
- Our backlog increased dramatically ...

### #Records harvested by Elements, vs. #Records published to the repository



#Records harvested by Elements
Processed into DRO

### Why a backlog?

### No guidelines

- No clear guidelines to assist with prioritising material
- No guidelines on who owns the workflow, and who is responsible for what, as well as timelines

Administratively cumbersome

- Everything is nominated for HERDC > everything goes through Faculty
- Letters to the editor receive the same attention/checking as peer reviewed articles
- Faculty view all pdfs to verify quality, regardless of publication source
- Liaison with the researcher to obtain pdfs leads to hold ups

Inefficient workflows and metadata feeds

Work from a priority list, no guaranteed turnaround time
 ERA / Staff promotions / New researchers / 2011-2017 publications

### Impact on researchers

- > Staff profiles out of date
- > Delays in making Deakin research discoverable
- > Impacts on citation rates
- > Complicated process that researchers
  find confusing

I am new to Deakin ... this publications workflow is so complicated!! I have accepted my publication – why is it not on my Staff profile!?!

0

đ



# Solutions and improvements

### **Our values & assumptions**

- There is value in retaining Faculty expertise in the pubs workflow
- Research publications need to be vetted and classified as they come thru Elements
- Non reportable outputs don't need the same high level of processing as reportable publications
- There needs to be a "One Deakin" approach toward the processing of publications
- The repository is important and is valued:
  - to Preserve research
  - to make it Discoverable
  - to make research Openly Accessible (to increase citation rates)

### Consultative workshops, 2016-17

### • Combined Library, Faculty, and Research Office workshops

- To provide a forum to voice opinions and concerns
- To map out and identify variations between the Faculties in their approach
- Workshops including the Library, Deakin Research and Leonie Hayes from Symplectic
  - To review metadata-based workflow and systems triggers in the Elements > DRO > Research Master data feeds
  - To re-engineer workflows and data feeds
  - To review harvested metadata and determine ways we can automate our processes whilst ensuring quality

# New "Smart ingest" approach to harvested metadata

- Journal articles should be accepted as C1 or C1.1, and meet the definition of research without additional checking if:
  - The publication was found by Scopus/Web of Science
  - The publication is of type "article" in SCOPUS/Web of Science
  - The publication is more than 2 pages long
  - The parent journal is in SSCI/SCIE or in SCImago Q1 (as per Where Should I Publish)
  - The title does not include the words "Review", "Editorial", "Case Study" or other known stop words likely to indicate the item does meet the definition of research.

- A spreadsheet of claimed journal articles in Elements is generated nightly and output to a shared Library/Faculty wiki site
- The report is generated by a custom PHP script running on DRO
- The script queries all the journal records in Elements, does an ISSN match against a pre-loaded dataset that includes various quality matrices (e.g: SCImago journal score, quartile, impact factor, number of pages, etc.

F	ILE	HOME	IN	SERT	PAGE LAY	OUT	FORMU	LAS	DATA	REVIEW	VIE	W D	EVELOPE	R AC	ROBAT											II.	
K:	.6	Ŧ	: ]	K 🗸	f <sub>x</sub>	2655487	00011																				
	А	в	С	D	E	F	G	Н	I	J	к	L	М	N	0	Р	Q	R	S	Т	U	V	V	×	Y	Z	AA A
1	DRO p	u Elemen	HERD	C Title	Publica	t Publicat	Start p	a End pa	No. of p	Scopus	WOS id	Publicat	WSIP IS	5 Journal	ARWU	Impact :	ADBC ra	Scimago	Top qu	WSIP re	Author	Author	Faculty	School			
2	DU:3007	Chttps://s	https:/	/sy Hierarch	ni 2010	Journal Ar	ticle		1		2.1E+11	657727	0065-77	2 Abstracts	SCIE					Recomme	Matthew	Barnett	GTP Rese	Institute	for Fronti	er Materials	;
3	DU:3007	6 https://s	https:/	/sy Improvi	ng 2010	Journal Ar	ticle		1		2.1E+11	657727	0065-77	2 Abstracts	SCIE					Recomme	Kieran	Lim	Faculty o	f School o	f Life and E	nvironment	al Sciences
4	DU:3007	Chttps://s	https:/	/sy The resp	o 2009	Journal Ar	3768	3 3774		2-s2.0-69	54911821	1E+07	1000-09	3 Acta Ecolo	ogica Sinio	а		0.172	Q4		Margaret	Ackland	Faculty o	f School o	f Life and E	nvironment	al Sciences
5	DU:3007	Ehttps://s	https:/	/sy Solute c	lu 2010	Journal A	4923	4939	17	2-s2.0-77	2.8E+11	1.4E+07	1359-64	5 Acta Mate	SCIE	5.058		3.683	Q1	Highly Red	Ross	Marceau	GTP Rese	Institute	for Fronti	er Materials	
6	DU:3007	Ehttps://s	https:/	/sy Evolutio	n 2010	Journal A	1799	5 1805	11	2-s2.0-74	2.7E+11	1.4E+07	1359-64	5 Acta Mate	SCIE	5.058		3.683	Q1	Highly Red	Ross	Marceau	GTP Rese	Institute	for Fronti	er Materials	;
7	DU:3007	1 https://s	https:/	/sy Detache	ed 2010	Journal Ar	202	2 206		2-s2.0-77	2.8E+11	1.6E+07	1601-52	1 Acta Neu	SCIE	0.76		0.421	Q3	Recomme	Michael	Berk	Faculty o	f School o	fMedicine		
8	DU:3007	1 https://s	https:/	/sy Pharma	cc 2010	Journal Ar	99	5 97		2-s2.0-77	2.8E+11	1.6E+07	1601-52	1 Acta Neu	SCIE	0.76		0.421	Q3	Recomme	Michael	Berk	Faculty o	f School o	fMedicine		
9	DU:3007	1 https://s	https:/	/sy Psychos	o 2009	Journal Ar	275	5 284		2-s2.0-70	2.7E+11	1.6E+07	1601-52	1 Acta Neu	SCIE	0.76		0.421	Q3	Recomme	Michael	Berk	Faculty o	f School o	fMedicine		
10	DU:3007	1 https://s	https:/	/sy Psychos	o 2009	Journal Ar	275	5 284		2-s2.0-70	2.7E+11	1.6E+07	1601-52	1 Acta Neu	SCIE	0.76		0.421	Q3	Recomme	Lesley	Berk	Faculty o	f School o	fMedicine		
11	DU:3007	1 https://s	https:/	/sy The con	ce 2009	Journal Ar	272	2 274		2-s2.0-70	2.7E+11	1.6E+07	1601-52	1 Acta Neu	SCIE	0.76		0.421	Q3	Recomme	Michael	Berk	Faculty o	School o	fMedicine		
12	DU:3007	1 https://s	https:/	/sy Validity	o 2010	Journal Ar	100	101		2-s2.0-77	2.8E+11	1.6E+07	1601-52	1 Acta Neu	SCIE	0.76		0.421	Q3	Recomme	Michael	Berk	Faculty o	School o	fMedicine		
13	DU:3007	1 https://s	https:/	/sy Writing	le 2009	Journal A	314	316		2-s2.0-70	2.7E+11	1.6E+07	1601-52	1 Acta Neu	SCIE	0.76		0.421	Q3	Recomme	Michael	Berk	Faculty o	f School o	fMedicine		

- Faculty staff download the spreadsheet and apply filters to identify journal articles that automatically qualify as C1s
- The spreadsheet has been greeted with enthusiasm by most Faculty staff
- It means they don't have to view every publication that's coming through, and can confidently assign a C1 classification to articles that meet the criteria
- Processing time has been doubled using the new techniques

# Next steps: working with Symplectic to further automate our workflow

- The spreadsheet is useful for automatically identifying C1s, but the Elements interface requires Faculty to process records individually
- Working on utilising the Elements API to automatically update records identified as C1s
- Can we also implement some of these automatic filters when the record is first harvested?
- Can we remove the need for researchers to complete the deposit process for every publication?

### Other solutions and improvements

### New, more efficient metadata feeds



A nightly feed fed metadata from **DRO** 

*Elements has become the source of publication truth* 

Metadata fed directly from Elements to Research office and staff profiles

### New supporting guidelines and policies







- High level document approved by Academic Board in March 2017
- Defines Roles and responsibilities of Researchers, Faculty, the Library and Deakin Research
- Splits research publications into Reportable and Non reportable material

#### **Research Guidance Sheet**

- An official Research Guidance Sheet authored by Deakin Research
- Aimed at Faculty administrators
- Advocates a consistent 'One Deakin' approach to processing publications
- Officially sanctions the 'Smart ingest' principles to identify C1s

#### **Research Repository Guideline**

 Outlines the role and purpose of the repository, emphasising preservation, discovery, and making research open access

DEAKIN UNIVERSITY

GUIDELINE

Information Property little

Crog Anikano

and Transfer Committee Class

Red And Thilesonity December 1

RESEARCH REPOSITORY

 Specifies what research outputs DRO collects (along with other material collected by the Library)

## Revamped wiki space to provide assistance and guidance for researchers

### Provides researchers with info about: The publication workflow, ar

- The publication workflow, and roles and responsibilities
- How to optimise their profile to facilitate efficient processing
- Linking their ORCID and Elements accounts
- How to use Elements
- FAQs

utions and improvements

- Tools and resources for download
- How to import and export
- Contact us
- Feedback
- Policies and guidelines, etc.



Before you start Publication fundamentals | The researcher's role in the publications workflow | Using DeakinSync to track your publications



<u>Getting started</u> <u>Optimising your Elements profile</u> | <u>Navigating Elements</u> | <u>What is Elements?</u> | <u>How to link your Elements / ORCID accounts</u> | <u>Delegating access to your account</u>



Depositing your research
Claiming and reviewing a publication
Completing the deposit process
Creating a new manual publication



Open Access Making your research open access in DRO



Importing and exporting publications Import publications into your Elements profile Export publications from Elements



Bibliometric data Bibliometric data in DRO | ReMAS (external link) | WoS, Scopus and h-Index data in Elements



FAQ, tools and resources Publications FAQ Tools and resources available for download



Policies and guidelines





### *Toolkit of downloadable infographs, e.g. 'How to Optimise your profile'*



### 'Progress of your publications' tile in DeakinSync



### Next steps

- Focus on Scopus ID cleanup and uptake
- Work with Symplectic:
  - to develop the Elements API to automatically update articles identified as C1s
  - Investigate further improvements to support automated processing
- Work with Faculty:
  - to ensure they are happy and confident with the new approach
  - to target the non-reportable records that are stuck in the queue



### michelle.watson@deakin.edu.au