

Database Preservation at the National Archives of Australia



What I'll talk about today:

- History of digital records at NAA
- Existing approaches to database transfer and preservation – 2 examples
- Database Preservation Project

A brief history of digital preservation at NAA





- Transferred on physical media
- Magnetic tape/floppy disks

c. 1970-1996

- Often printed out
- Poor description
- Obsolescence

1996-2000

Distributed custody

- · Agencies retain custody
- Management regime developed by NAA
- Period of standards development





Digital Preservation Project

- Bespoke DP software platform
- Commissioned in 2007
- Transfers received from 2000
- No specific approach to database transfer

2001-2017



2017-2020

Digital Archives Taskforce

- Dedicated Team (12 FTE)
- Tender for Integrated Archival Management System (IAMS)
- Digital Capability Strategy
- Policy updates
- Archival Control Model update





Integrated Archival Management System project

- Preservica acquired and commissioned
- Upgrades to existing inhouse platforms – archival management system, digitisation

2020 onwards

Digital archiving projects

- Digital Capability Strategy
- Research Framework & Plan: Emulation, Database preservation, Access, Description – support operational areas



National Archives of Australia



```
D1ÜÄÖâãÉCÁŐĎØĎ÷ØÓĎØĎ÷ØÓ Dä D…7×
                                                                                                               NŸÆÖÙ@ãÈÂ@ÈÂÁÙ@Âäã@âÈÖæÂÄ@A@ŐÂã@ÓÖââ@ÖŐ@ÉŐÃÖÔÂ@ÖÆ@「ñ÷kòòôKøóK@d@@@@@@@
                                     O ÄÖâãÉÇÁŐÒØÑĎØÓÒØÑĎØÓ 🏻 💯
                                                                                                                                                                                                                                                                                                         NŸÃÖÂÃ
                                                                     DI+ÄÖÄÄÉÇÁŐÑÓÖÔØÔÑÓĎÔØÔ DO DI.ÂÄÉ@ÜÄXÖÙÄØÖÆ@ÈÄÄÉŐÇ@ÄÉÇÈÄÄÄ@ÜÀÈÖÖŐÄ@ÕAŐÄ}Â@ÖÖÄÖÜĞÂÀÈÉÄÓÅ@@@@@@@@@@
œaååòóèœâåàùőéőçãœôùœâxùöäóåœåáåäöóxóéâèÅázœœœœœœœœœœœœœœ uíbãöããéçàőòŏñøôòổðñøổ u% uía[ñððœœmáxxùöç]œœœœáèùéããóàà@áóäåæáóààæáðàà
OÍXÃÖÂÃÉCÁŐÒÕÕÑØÔÒÕÑØÔ O% ÓĪWÙÂŐÃÂÄK@@@@@@@@@@@@@@@@@@@
DÔVĂÖÂÃÉCÁŐÓÑĎÑØÔÓÑĎŇØÔ D% DÔUÃÈÅ@ÉŐÆÖÙ
              DGPÄÖÄÄÉCÁŐÒ÷ĎÒØÔÒ÷ĎÒØÔ D* DGOMÄÈÄŐAÈTGÉŐQÙÁÓÁÄÉÖŐQÃŐQÁÄÂXÄÄÄÄÄQÖÆQÖŐÄÉÄÈTÄQÖÖÄÉÖCÄÉŐCAKGQÉŐGQQQQQQQQ
                                                                                                                                                                                                                                                               IGOÃÖâãÉCÁŐÒ÷ÖÒØÔÒ÷ÖÒØÔ
                                                  999999999999947ÚAÓ98ÁAKKÖÚX9ÁÁBÁĞÖÁÐÉBÁÁGÐÖÖÖÖÐGÉÐÉBÁKÁSÁKÐÁKÁTOTU "IGEÓÐÓÐOÐÐÐÐÐOÐÐÐÐOÐÓÐÁGÁKÁBÁKÁBÁKÁTÍ
                                                                                       DO DO PROPOSO PO PROPORTO DO PROPORTO DA CARA PROPORTO POR PORTO POR PROPORTO POR PROPORTO POR PROPORTO POR PORTO POR PROPORTO POR PORTO POR PORTO POR PORTO POR PORTO PORTO POR PORTO POR PORTO PORTO POR PORTO POR PORTO PORTO POR PORTO PO
                                                                                                                            OHCÁÖÂÃÉÇÁŐÒ÷ÖÒØĜÒ÷ÖÒØĜ O O OHBÁÙÁGÃÉAŐĞÂÃÚÉXXÂÄGÓÁÂÂÉŐÇĞÃÉAŐĞAÂGÁÖÁÂĞĞÀÓÓÁÂGÂÈÁÓÓÁKG
 K@@ÆÉŐÁÓÓÈk@ÁÂÂÂÃÃÃÂBÁÚÅ@ÃÖÖÄBÄÖÖÄBÄÖÖXÁŐÉÂÂ@@@@@@@
ÄÈÁÂÄÄGÂÉCGÙÁŐÃGXÙÖÄÄÄÉŐCGXÙÖXÂÙÃÉÅÂGÆÈÉÄÈGEÃGÃÖÓÄGÆÉÃÈÉŐGGGGGGGGGG
                                                                                                                                                                 OH9ĂÖÂÃÉÇÁŐÒ÷ÖÒØÔÒ÷ÖÒØÔ O* OH8ÃÈ®XÂÙÉŐÄK®®ÃÈ®ĂÁXÉÃÁÓ®XÙÖÆÉîÆ
□-ÅÃÖâãÉCÁŐÑÖðÖØÔÑÖðŐØÔ □º □-ÁÖ
                                □-'ĂÖâãÉCÁŐÑÖŐŐŐŐŐŐŐŐŐ □* □- ŐÖÙÀ@âäấ×ÀĂãÁÄ@CÙÖä×@ĂÖÔ×ÀŐÉÁâ@ÅÖãÉ@ÉŐ@ŐKâKæKk@ÉŐãÀÜÂãÁãÁK@@@@@@@@@@
                                                                     * D®>K@@@@AÄãÉâÉãÉÂÂ@ÉŐÄÓÄÄÉŐÇ@ÉŐãÙÀ`ÇÙÖäX@AŐÄ@ÉŐÃÂÙŐÁãÉÖŐAÓ@@@@@@@@@@@@@@@@
                                                                                                                                                                                   □®»ÃÖâãÉÇÁŐÑÖÖŐØÔÑÖÖŐØÔ □º □®ÔK@@@@ãÁŐâÁÃãÉÖŐâK@@MÔÖâã@
□ TAĂÖâãÉCÁŐÑ÷ðŐØÔÑ÷ðŐØÔ □ º □ T∠MÁ]@@ĂÁXÉ
              □ ÃÖÂÃÉCÁŐÑ÷ÖŐØÔÑ÷ÖŐØÔ □ º □ NK@@@@ÂÄÂŎÉŠÃÄÄ@ÉØ@DÀÂXÂÃéÆ@ÃæÖ@OÖÖĂÉÂÈ`ÓÉÂÔ@ÃÖÔXÁŐÉAÂK@@@@@@@@@@
                                                                                                                                                                                                                                                               ¶ÄÖâãÉÇÁŐñ÷ðőøôñ÷ðőøô
                                                  POPOSOBIALA DE CAPACIÓN DE CAP
                                                                                                                         á⊃ÄÖÖĞÜĞÚÂ×®ÉÉÊŤĞÜ ÷□ ÒœĞñĞñŏñŏñŏñŏñŏñŏā
ùäâã@ÃÖÔ×ÁŐÉÂâ@ÉŐåÖÓåÄÄ@@@@@@@@@ hoozøo€o
+ÁÖÂÃÉCÁŐÑÔÑĎØÓÑÔÑĎØÓ U÷ UĎÌÖÖÄCÂÄ@ÁÃ@ÔÂÓÂÖÄÜÖÄDÉÖ@ÑÙ÷ÙK@ÉÖ@ÃÊÂ@ÙÁÃÃÜÖK@ÃÊA@XÂÙÂÖÖ@@@@@@@@@@@@
                                                                                                                                                                                                                                            _OðiÃÖâãÉCÁŐñðñðøóñôñðøó — OrÓÃ
                                DZGĂÖÂÃĖÇĀŐÑÔÑÔØÓŇÔÑÔØÓ D DŽFŐÂÇÖÃÉÄÃÃÄØÄÃ@ŒÂÊÂ@ÃÔÉ@ÈÁæÃĚÖŮŐ@ÜÖÄÄK@ÖŐ@ÒŐÑØÑ@Ţ÷ÖØKŎŎ@@@@@@@@@@@@@@@
                                                                    apágagagaga káákódóðó gó dío ag vila kákót ag vila kákót ag vila kákót ag vila kákákót ag vila káta kákót ag vi
                                                                                                          □ZnÃÖâãÉCÀŎÑÔŘÒøÓÑÔďÓĎÓ □ □ZmÃÖŐãÀÉŐÂ@ãÁC@ÙÀÆäŐÄGÁÈÄØäÀ@ÆÖÜ@ŢÖÒ÷KÓŌ@ÄÁãAãGÒ÷ŌøøÑ@AÓÄ@Œ
DÍÏÃÖÂÃÉÇÃŐÒŐŐNØÔÒŐNØÔ DŚ DÍÍKGGGGGGGGGGGGGGGGGGÑÖÈŐGÆÚÁŐÃÉÂGÖÌÄÖŐÖÈÄÂGGG
ODDODODODODO DA BABO DO DO DA COMPAÑA DA CARRA C
                                                                                                                                                                               DÌAÄÖÂãÉÇAŐÒÕÕÑØÔÒÕĎÑØÔ 0% DÌÄæAÙAØAÄAÔAŐÃŒÃÉAÃŒÂXÜÖÄÓA
DÌBÃÖÂÃÉCÁŐÒÕÕÑØÔÒÕĎÑØÔ D% DÌÞÆÙÖÔ@ÃÈÂ@
              DÌÙÄÖÄÄÄÉÇÁŐÒŐŐÑØÔÒŐĎŐØÔ U% DÌØÆÙÖŐØÈÉÂGÁÁÙŎÉŐÇÂKŐŐØÙÄÁÖÙÄÄGÖÆŘÉÁGXÁÙÉÖÄĞÄÈÁÄŐØÉÄÈÓÓÓÂOØOØØØØØØØØØØ
 ùk@Èå@ÈàÄ@æÖùÒàÄ@àã@A@@@@@@@@@@@ I1$ÃÖâãÉCÁŐÒŐÖñøôòÖñøô I% II#âàùÈÖäâ@AÔXÓÖèâùâ@æAâ@XáÈÄ@æÈåŐ@ãÈÂ@ÑÖAæAâ@A6ÖÖA
 Aâ@A@XÁÉŐÃÅÛ@ÁŐÄ@ÄÖÄÒÅÙK@ÃÈÉÂ@XÂÛÂÖŐ@æÁÂÂ@@@@@@@@@
                                                                                                                             DÍIBÃÖÂÃÉCÁŐÒŐŐÑØÔÒŐŐÑØÔ I% DÍBÁ@}XÖŐ}@ÁÖÄØÈÁÄ@ÂÉŐÃÂ@ÛÁÃÄÜÓÁÄŒÃÖQÁŐCÓÁŐÄK@ÔÙ@ÂXÙ
 OÍCÃOÂÃÉCÂŐÒŐŐÑØÔÒŐŐÑØÔ 0% OÍBÙÂÂÉÄÂÄØAÃK®æÈÃŐ®äÂÉŐC®CÓÄÄÂÃŐŐÂ}
AGZAÚAGÚAGAAAAAAGAÖG F÷ŐKÖÖÖAGGGGGGGGGGGGGGG
 cãgôöőäÁèk@@ãèÂ@ãÂùô@öÆ@ãèÂgóÂŐÄÉŐC@âÖäCÈã@æÁâ@@@@@@@@@@
                                                                                                                                         □b-{ñöðòpôañöðòpô{@ãÁçaðőôðaò`óκ@@@@@@@@@@@@@@@@@@@@@@@@@@@
                                                                                                                                                                                □þ-ĂÖâãÉÇÁŐÒØĎŐØÔÒØĎŐØÔ □* □þ-ÖŐ@ÑÖĎÒØÔ@ÑÖÈŐ@ÂäçãÖŐ@ÁŐÄ
Db±ÃÖÂÃÉCÁŐÒØĎŐØĎŐØĎŐØÔ Dº Db°ÃÉA@DÃÉÉŐÂÂADØĂÖŐŐÄÃÃÉÖŐÂQÉÖ@EÖŐC@ÒÖÖÇ@AŐÄĞAÖAÂE@æADA@ŒÊÂĎ@@@@@@@@
                                                                                        Upoãoâã£caóògôógôògôògôoo u~ UpoãáócQ&áôéóéââQáóÁQôùQæöőcQQáQóQAQÓAQÁAAAQôùQâacãoóQãè
őåè@ãèùöäcè@ãèå@óéåô`ãåŏ`@@@@@@@@
 9999999999990584044x9666]9003x94564489494460
                                                                                                                            ðñ÷ñaðððñ÷ñ{@@ãAcaðòópaỗñpK@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@ [SãÃÖâãÉCÁŐñóðôpôñóðôpô [] [SãA@ñùöö@ÃAãAÓAÙAŐ@ÁÙäÉâAÙďæAâ@xäÙÆ
\acute{o}\ddot{o}\acute{o}\acute{o}\acute{o} \ddot{o} \ddot{o}
                                                                                                                                                                                                      OŠÙÄÖâãÉÇÁŐÑÓÔÔØÔÑÓÔÔØÔ 0] OŠØMÂÅÅ®XÇ®ÓÑØ®ÆÖÙ®Æ
1ÃÖÂÃÉÇÃŐÑÓĎÔØÔÑÓĎÔØÔ 0 0 0 SÉÆÂÙÂŒRÁÉŽÉŐÇŒÆÜÐŰĞÂŐÙÒŒÂŐÙÒŒÂĞÜÃÖĞÂĞÂŒÂÂÊÂÂŒXÓAČŐÄÄŒÃÖŒÇÖŒŒŒŒŒŒŒŒ
                                                                                                                                                                                                                                            OŠĖÅÖåãÉÇÅŐnóðôøônóðôøô O] OŠ÷Æ
                                 $A&v@AA&@A@$£&@EA$$&£A$$@£$@A&$A£&$@£A&&&&&&&A&&@@@@@@@@
                                                                                                                                               lacksymbol{	ilde{n}} lac
```

A brief history of Digital Preservation at NAA





- Transferred on physical media
- Magnetic tape/floppy disks

c. 1970-1996

- Often printed out
- Poor description
- Obsolescence

1996-2000

Distributed custody

- · Agencies retain custody
- Management regime developed by NAA
- Period of standards development





Digital Preservation Project

- Bespoke DP software platform
- Commissioned in 2007
- Transfers received from 2000
- No specific approach to database transfer

2001-2017



2017-2020

Digital Archives Taskforce

- Dedicated Team (12 FTE)
- Tender for Integrated Archival Management System (IAMS)
- Digital Capability Strategy
- Policy updates
- Archival Control Model update





Integrated Archival Management System project

- Preservica acquired and commissioned
- Upgrades to existing inhouse platforms – archival management system, digitisation

2020 onwards

Digital archiving projects

- Digital Capability Strategy
- Research Framework & Plan: Emulation, Database preservation, Access, Description – support operational areas



National Archives of Australia

Current approach at NAA

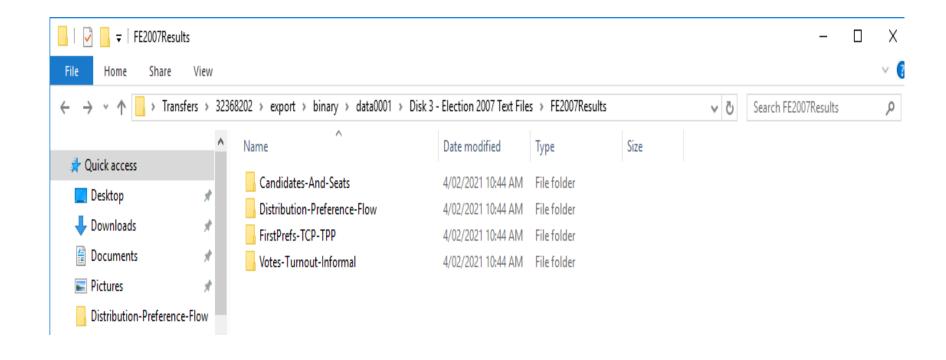
- Not many database transfers received
- Transfers are ad hoc no standard approach
- We've set a very low bar for agencies (keep it simple)
- Data exports in an open format plain text, CSV
- Transfer very little technical or structural metadata
- Haven't always asked for data dictionaries or other technical documentation
- Consequently, we have very little information about how these systems were managed

Examples from the NAA collection: 1. 2007 Federal Election Results

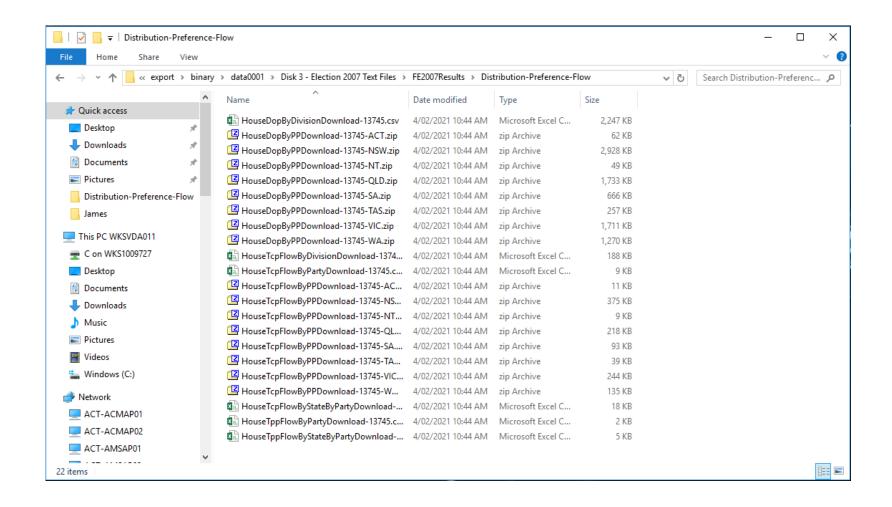
Federal Election Results – Disposal Class for Election Results

Disposal Action Description of Records Entry Retain as national Class no. Final versions of significant reports and returns on the conduct of elections, ballots and referendums. Includes: archives 11729 official reports on Federal elections presented by the Electoral Commissioner following an election · summary reports on the conduct of referendums official results of elections or referendum in statistical tables divisional returning officer returns election journals.

Transferred Data



Transferred Data



Transferred Data

☐ 5 · ♂ · ₹ HouseCandidatesDownload-13745.csv - Excel																	
File Ho	ome Insert	Page Layout For	mulas Data R	eview View Deve	loper Conten	t Manager 👂 Tell me	what you want to do										
📇 🐰 Cut	Calibri	- 11	A A = =	- ₩vrap Tex	t Gene	el v B		lormal	Bad	Good	Neutral		Calculatio			× 🛗	∑ AutoS
Cop	v *						#							<u>"</u> -			▼ Fill +
Paste Forn	nat Painter B I	<u>u</u> - 🗄 - 💆	- <u>A</u> - = = =	€ ₹ Merge &	Center - \$ -		litional Format as	heck Cell	Explanato	ry Input	Linked	Cell	Note		Insert De	lete Format	
Clipboar		Font	G .	Alignment	G.	Number 5	itting rubic			Styles					C	ells	
		. 1															
A1	▼ : × ✓	<i>f</i> _x 2007 Fe	deral Election House	of Representatives C	andidates [Ever	t:13745 Phase:FinalRes	ults Generated:200	08-09-03T12:42	2:03 Cycle:34	9e38e7-5eda-4bc7-a	4a0-805bff5	9620f Cre	eated:2008-	·08-13T17:	22:12 Envir	onment:TES	T Site:Pri
A	В	С	D	Е	F	G	н		1	J	К	L	М	N	0	Р	Q
1 2007 Fede	ral Election House					:2008-09-03T12:42:03 Cy		a-4bc7-a4a0-8								/TR04 Versi	
2 StateAb		DivisionNm	PartyAb		CandidateID	Surname	GivenNm	Elected		HistoricElected				_ ·			
3 NSW	151	Warringah	LP	Liberal	1	3029 ABBOTT	Tony	Υ		Υ							
4 QLD	162	Forde	FFP	Family First	1	7896 ABRAHAMSON	Iona	N		N							
5 TAS	196	Lyons	ALP	Australian Labor Pa	1	7500 ADAMS	Dick	Y		Υ							
6 NSW	152	Wentworth	FFP	Family First	1	7721 ADAMS	James	N		N							
7 NSW	132	Mackellar	CCC	Climate Change Coa	1	3387 ADAMS	John	N		N							
8 VIC	210	Dunkley	GRN	Australian Greens	1	3367 ADAMS	Neale	N		N							
9 VIC	227	McMillan	FFP	Family First	1	7647 AESCHLIMANN	Terry	N		N							
0 NSW	121	Grayndler	ALP	Labor	1	3196 ALBANESE	Anthony	Y		Y							
1 VIC	227	McMillan	CEC	Citizens Electoral C	1	7571 ALBLAS	Theo Matthe	w N		N							
2 NSW	126	Hunter	CEC	Citizens Electoral C	1	3170 ALBURY	Daniel	N		N							
3 VIC	234	Wills	SAL	Socialist Alliance	1	7798 ALCORN	Zane	N		N							
14 VIC	208	Corio	FFP	Family First	1	7910 ALDERSON	Gordon	N		N							
5 NSW	134	Mitchell	CDP	Christian Democrat	1	3280 ALLEN	Darryl	N		N							
6 NSW	105	Bennelong	IND	Independent	1	7233 ALLEN	David L	N		N							
7 NSW	109	Calare	CLR	Country Labor	1	7483 ALLEN	Michael	N		N							
8 SA	191	Wakefield	GRN	The Greens	1	7991 ALLEN	Terry	N		N							
9 VIC	309	Gorton	FFP	Family First	1	7645 AMBERLEY	Scott	N		N							
0 NSW	149	Sydney	LP	Liberal	1	7918 ANDERSON	Georgina	N		N							
1 VIC	229	Menzies	LP	Liberal	1	7256 ANDREWS	Kevin	Υ		Υ							
2 NSW	113	Cowper	FFP	Family First	1	7722 ARAPI-NUNEZ	Flavia	N		N							
3 VIC	217	Hotham	LP	Liberal	1	7282 ARBOREA	Vince	N		N							
4 NSW	249	Paterson	ALP	Labor	1	8081 ARNEMAN	Jim	N		N							
5 ACT	102	Fraser	CEC	Citizens Electoral C	1	7311 ARNOLD	Jim	N		N							
6 NSW	142	Prospect	LP	Liberal	1	7843 ARTHUR	Lily	N		N							
7 VIC	210	Dunkley	FFP	Family First	1	7663 ASHDOWN	Steven	N		N							
28 QLD	167	Kennedy	DEM	Democrats	1	3346 ASPLIN	Nigel	N		N							
9 NSW	137	North Sydney	IND	Independent	1	3458 AUSSIE-STONE	Marcus	N		N							
0 TAS		Denison	SAL	Socialist Alliance		7893 AUSTIN	Susan	N		N							
1 NSW		Page	CDP	Christian Democrat		3372 AVASALU	Rhonda Joy	N		N							
2 VIC		Aston	DEM	Australian Democra		7950 AZA	Rachal	N		N							
33 NSW		North Sydney	CDP	Christian Democrat		3299 BAALBERGEN	Arie	N		N							
34 QLD		Griffith	GRN	The Greens		3139 BACH	Willy	N		N							
S5 VIC		Higgins	FFP	Family First		7639 BADWAL	Penny	N		N							
6 VIC		McEwen	LP	Liberal		7265 BAILEY	Fran	Υ		Υ							
7 VIC		Dunkley	DEM	Australian Democra		7958 BAILEY	Karen	N		N							
88 VIC		Goldstein	DEM	Australian Democra		7954 BAILEY	Michael	N		N							

Series Registration for A13777: 2007 Federal Election Results

C							
Series details for: A1377							
Series number	A13777						
Title	Federal Election results						
Accumulation dates	1901 -						
Contents dates	2007 -						
Items in this series on RecordSearch	4 All items from this series are entered on RecordSearch.						
Agency/person recording	21 Feb 1984 - CA 3854, Australian Electoral Commission, Central Office						
Agency/person controlling	CA 3854, Australian Electoral Commission, Central Office						
Quantity and location	0.2 metres held in ACT						
System of arrangement/ control	Single number series						
Predominant physical format	ELECTRONIC RECORDS						
Series note	Summary heading						
	Function and purpose						
	Related legislation						
	Using the series						
	Language of material						
	Physical characteristics						
	System of arrangement and control						
	Relationships with other records						
	Finding aids						
	Access conditions						
	Series history						
	Provenance						
	Immediate source of acquisition						
	Custodial history						
	Quantity in agency custody						
	Disposal history						
	Publication note						
	Additional information						
	End notes Sources						
	Sources						
Visibility & availability	(Less) 73 . All items from the series are on RecordSearch						
indicator	20 km 2000						
Date registered	29 Jun 2009						
Date modified	24 Jul 2014						

Item Registration for A13777

Item details for: A13777,	3 Issue to research centre Request copy
Title	Federal Election Results - 2007 - text files
Contents date range	2007 - 2007
Series number	A13777 Click to see which government agency or person created this item.
Control symbol	3
Citation	NAA: A13777, 3
Item ID	32368202
Location	Canberra
Access status	Not yet examined
Physical format	ELECTRONIC RECORDS (allocated at series level)
Series accession number	2009/00555361
Records authority class number	11729
Date registered	24 Jul 2014
Date modified	24 Jul 2014

Issues

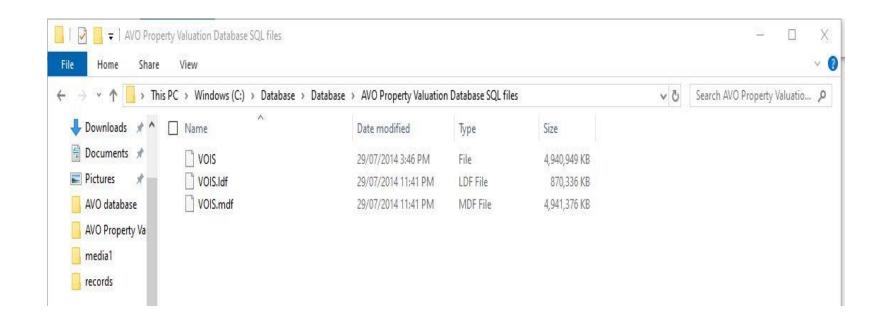
- Description
- Technical documentation
- Raw data format
- You can probably think of other issues

Examples from the NAA collection:

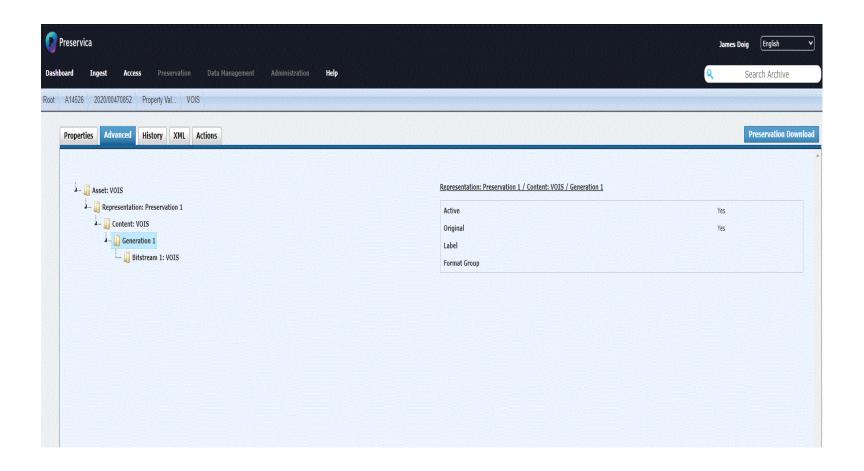
2. Australian Valuation Office Database (2014)

National Archives of Australia

Native SQL files



Preservica file format identification



XML files of database tables

→ ↑	is PC > Windows (C:) > Database > Database	 AVO Property Valuation 	n Database XML Files	
Quick access	☐ Name	Date modified	Туре	Size
7-1,7-100,00-014-0-0.000	dbo_ACTIVITY.xml	24/07/2014 12:12 PM	XML Document	11 KB
Desktop 💉	dbo_ACTIVITY.xsd	24/07/2014 12:12 PM	XSD File	12 KB
▶ Downloads 🖈	☑ e dbo_BM_LOCALITY_CODES.xml	24/07/2014 12:12 PM	XML Document	41 KB
Documents 🖈	dbo_BM_LOCALITY_CODES.xsd	24/07/2014 12:12 PM	XSD File	7 KB
📰 Pictures 🧳	dbo_BM_LOCALITY_POSTCODES.xml	24/07/2014 12:12 PM	XML Document	549 KB
AVO database	dbo_BM_LOCALITY_POSTCODES.xsd	24/07/2014 12:12 PM	XSD File	8 KB
AVO Property Valua	dbo_BM_POSTCODE_VALUES.xml	24/07/2014 12:13 PM	XML Document	15,596 KB
AVO Property Valua	dbo_BM_POSTCODE_VALUES.xsd	24/07/2014 12:13 PM	XSD File	9 KB
media1	dbo_BM_PROJECTS.xml	24/07/2014 12:15 PM	XML Document	13 KB
inedia i	dbo_BM_PROJECTS.xsd	24/07/2014 12:15 PM	XSD File	7 KB
OneDrive	dbo_BM_VALUES.xml	24/07/2014 12:15 PM	XML Document	1,733 KB
This PC	dbo_BM_VALUES.xsd	24/07/2014 12:15 PM	XSD File	11 KB
RECURSIONS	dbo_COUNTRY_CODE.xml	24/07/2014 12:16 PM	XML Document	29 KB
3D Objects	dbo_COUNTRY_CODE.xsd	24/07/2014 12:16 PM	XSD File	5 KB
Desktop	dbo_CUSTOMER.xml	24/07/2014 12:16 PM	XML Document	87 KB
Documents	dbo_CUSTOMER,xsd	24/07/2014 12:16 PM	XSD File	14 KB
Downloads	dbo_CUSTOMER_CATEGORY.xml	24/07/2014 12:16 PM	XML Document	1 KB
Music	dbo_CUSTOMER_CATEGORY.xsd	24/07/2014 12:16 PM	XSD File	5 KB
■ Pictures	dbo_CUSTOMER_PROGRAMS.xml	24/07/2014 12:17 PM	XML Document	5 KB
Videos	dbo_CUSTOMER_PROGRAMS.xsd	24/07/2014 12:17 PM	XSD File	5 KB
	dbo_CUSTOMER_SERVICE_CENTRES.xml	24/07/2014 12:17 PM	XML Document	178 KB
Windows (C:)	dbo_CUSTOMER_SERVICE_CENTRES.xsd	24/07/2014 12:17 PM	XSD File	8 KB
Network	dbo_DATA_EXCHANGE.xml	24/07/2014 12:25 PM	XML Document	436,239 KB
	dbo_DATA_EXCHANGE.xsd	24/07/2014 12:17 PM	XSD File	90 KB
	dbo_DATA_QUALITY.xml	24/07/2014 12:27 PM	XML Document	1 KB
	dbo_DATA_QUALITY.xsd	24/07/2014 12:27 PM	XSD File	5 KB
	dbo_DEFENCE_ESTATE_REVALUATION	24/07/2014 12:27 PM	XML Document	434 KB
	dbo_DEFENCE_ESTATE_REVALUATION	24/07/2014 12:27 PM	XSD File	61 KB

XML file of a database table

```
o ×
        C:\Database\Database\AVO Property Valuation Database XML Files\dbo_BM_LOCALITY_CODES.xml
                                                                                                       → C Search...
C:\Database\Database\AVO... ×
  <?xml version="1.0" encoding="UTF-8"?>
- <dataroot generated="2014-07-24T12:12:36" xsi:noNamespaceSchemaLocation="dbo_BM_LOCALITY_CODES.xsd" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:od="urn:schemas-microsoft-com:officedata">
   - <dbo_BM_LOCALITY_CODES>
        <STATE>ACT</STATE>
        <LOCALITY_CODE>101</LOCALITY_CODE>
        <LOCALITY_NAME>Gungahlin & Hall</LOCALITY_NAME>
     </dbo_BM_LOCALITY_CODES>
    - <dbo BM LOCALITY CODES>
        <STATE>ACT</STATE>
        <LOCALITY_CODE>102</LOCALITY_CODE>
        <LOCALITY_NAME>Belconnen</LOCALITY_NAME>
     </dbo_BM_LOCALITY_CODES>
    - <dbo_BM_LOCALITY_CODES>
        <STATE>ACT</STATE>
        <LOCALITY_CODE>103</LOCALITY_CODE>
        <LOCALITY_NAME>Inner North Canberra</LOCALITY_NAME>
     </dbo_BM_LOCALITY_CODES>
    - <dbo_BM_LOCALITY_CODES>
        <STATE>ACT</STATE>
        <LOCALITY CODE>104</LOCALITY CODE>
        <LOCALITY_NAME>Weston Creek & Woden</LOCALITY_NAME>
     </dbo_BM_LOCALITY_CODES>
    - <dbo_BM_LOCALITY_CODES>
        <STATE>ACT</STATE>
        <LOCALITY_CODE>105</LOCALITY_CODE>
        <LOCALITY NAME>Tuggeranong & Queanb</LOCALITY NAME>
     </dbo BM LOCALITY CODES>
    - <dbo_BM_LOCALITY_CODES>
        <STATE>ACT</STATE>
        <LOCALITY_CODE>106</LOCALITY_CODE>
        <LOCALITY_NAME>Inner South Canberra</LOCALITY_NAME>
     </dbo_BM_LOCALITY_CODES>
    - <dbo_BM_LOCALITY_CODES>
        <STATE>ACT</STATE>
        <LOCALITY_CODE>107</LOCALITY_CODE>
        <LOCALITY_NAME>Cooma</LOCALITY_NAME>
     </dbo_BM_LOCALITY_CODES>
      -dh- DM LOCALITY CODEC
```

AVO Data Dictionary

Australian-Valuation-Office-Data-**Dictionary**¶ * Contents Australian Valuation Office Data Dictionary Audience Purpose... Contact-Information.. The AVODD Introduction.. Logical-structure-of-Appendable-Tables ... Points-related-to-the-AVODD Bench-Mark-Tables..... Sample SQL and Joins for Main Tables Locating-foreign-keys-from-the-filterable-excel-version-of-the-tabulations. The PROPERTY_ASSET_VAL-table-and-related-simple-look-up-table-joins The BENCH-MARK-TABLES and related joins TIME_SPENT-table-and-related-joins 17¶ The Data Tables and Data Elements 18¶ ACCESS_LEVEL .. ACTIVITY..... AREA_POSTCODES 19¶ ATO_AVO_GATEKEEPERS.. 201 AVO_ADDRESS... 201 AVO_GROUP. AVO_SKILLS .. 22¶ BM_LOCALITY_CODES. BM_LOCALITY_POSTCODES . BM_LOCALITY_POSTCODES_OLD. BM_POSTCODE_VALUES.... BM_PROJECTS.. BM_PROPERTY_TYPES. BM_VALUES. BOOKING_CATEGORY BULK_JOB_CODE... BULK_LODGEMENT_DEFAULTS .. CLK_DATA... CLK_FEES.. CONFERENCE_BOOKING_DETAILS. CONFERENCE ROOMS... COUNTRY_CODE.

Australian-Valuation-Office-Data-Dictionary-(AVOOD)-Page-4-of-1699

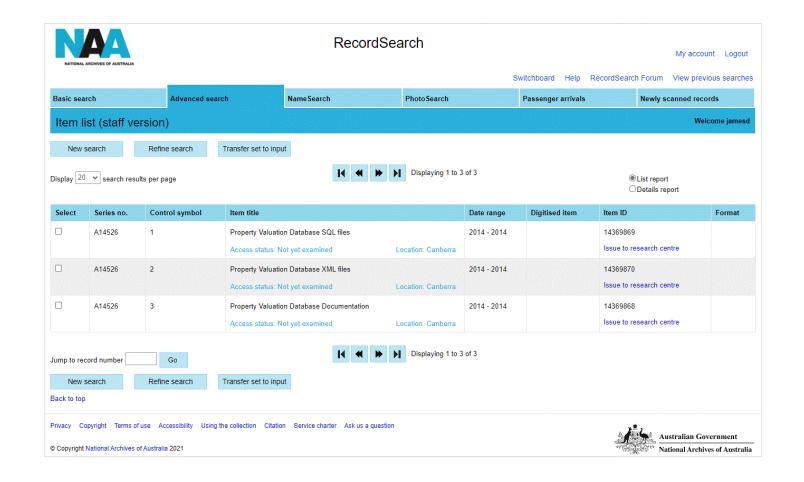
CUSTOMER_CATEGORY . CUSTOMER_PROGRAMS.

USTOMER_SERVICE_CENTRES	
arwinViewAllJobs	
arwinViewOnHandPE Other	
arwinViewOnHandPt_Other	
ATA EXCHANGE	
ATA_EXCHANGE_DATE_IMPORTED	
ATA_EXCHANGE_DATE_IMPORTED	
PataExchangeDisp10Oct13	
DEFAULT_TIME	
EFENCE_ESTATE_REVALUATION	
EFENCE ESTATE REVALUATION OLD	
EFENCE_PROJECT_STAGE	
EFENCE_REJECT_INFO	
EFENCE_REJECT_INFO_OLD	.81¶
TPROPERTIES	.81¶
VA_DATA	
VA_FEES	
VAL_SALES	
AT_LODGEMENT_DEFAULTS	
OURS_WORKED	
OUSE_CURTILAGE_TYPE	
NVOICEFEESSNAPSHOT	
OBS	
OBS_ASSET_TITLEOBS_ASSET_VAL	
OBS_ASSET_VAL_HISTORICS	
OBS_ASSET_VAL_HISTORICS	
OBS_PE_VAL	
OBSCOPY	
EAVE APPL	
IAISON INDICATOR	
OCATION	
OGIN LOG	1149
OOKUP_TABLES	114¶
IONTHLY_REPORT	115¶
IOU_TIMEFRAMES	127¶
IEW_AREA_POSTCODES	
OTICE_BOARD	
FFICE	
FFICER	
FFICER_GRADE	
FFICER_GROUP	
FFICER_LEVEL	
PFFICERBAK	
RIORITY	
ROJECT	
ROJECT_CUSTOMER_CODE	
ROPERTY	
ROPERTY ASSET VAL.	
ROPERTY CONDITION	
ROPERTY EXTERIOR	
ROPERTY_INTERIOR	
ROPERTY ROOF	
Australian-Valuation-Office-Data-Dictionary-(AV000)-Page-5-of-16	in .
	7.5

Series Registration for A14526

Series details for: A14526								
Series number	A14526							
Title	Australian Valuation Office Database, 2014							
Accumulation dates	14 Jan 2014 - 29 Jul 2014							
Contents dates	14 Jan 2014 - 29 Jul 2014							
Items in this series on	3							
Record Search	Click to see items listed on RecordSearch. Please contact the National Reference Service if you can't find the record you want as not all items from the series may be on RecordSearch.							
Agency/person recording	14 Jan 2014 - 30 Jun 2014 CA 7089, Australian Valuation Office, Central Office 01 Jul 2014 - 29 Jul 2014 CA 995, Australian Taxation Office, Head Office							
Agency/person controlling	01 Jul 2014 - CA 995, Australian Taxation Office, Head Office							
Quantity and location	0.1 metres held in ACT							
System of arrangement/ control	Single number imposed by the National Archives							
Range of control symbols	[1] to [3]							
Predominant physical format	ELECTRONIC RECORDS							
Series note	Function and purpose This Series is the legacy valuation database of the Australian Valuation Office called VOIS (Valuation Office Information System). The Series comprises three items. One item consists of technical database documentation including the data dictionary, which explains the structure of the database and the meaning of the data it contains. One item contains the raw SQL data in .mdf (SQL Server data file) and .ldf (SQL Server log file) format. The third item contains the data exported to XML format.							
	The VOIS database recorded all valuation data collected by the AVO. The VOIS database has 106 tables, several of which were in use in the decade up to 2014, when the database was decommissioned. Some of the VOIS tables were out of date or obsolete as of 2014, but were still retained in the database and are of historical interest. There are also some tables that were in sporadic use up to the year 2000 or earlier. The 106 tables comprise over 13,000,000 records.							
	Related legislation							
	Using the series							
	Language of material							
	Physical characteristics							
	System of arrangement and control Single running number imposed by the National Archives							
	Relationships with other records							
	Finding aids							
	Access conditions							
	Series history							
	Provenance							
	Immediate source of acquisition These electronic records were sent to the National Archives when the Australian Valuation Office (AVO) was closed on 30 June 2014.							
	Custodial history							
	Quantity in agency custody							
	Disposal history							
	Publication note							
	Additional information							
	End notes							
	Sources							
	(Less)							
Date registered	20 Nov 2020							
Date modified	20 Nov 2020							

Items in A14526



Problem is...

- Disconnect between development of disposal schedule and transfer
- We haven't decided what we need to preserve to ensure meaningful access in the future.
- The raw data may not be enough:
 - Native SQL format is software dependent
 - Flat file format (e.g. CSV) has limited usability
 - Need technical documentation to understand data
- Are there any characteristics of the database that we need to preserve to ensure meaningful accessibility and usability?
- There is not be a one-size-fits-all approach to database transfer and preservation.

Database Preservation Project: December 2020-July 2021

"to develop and agreed approach to the appraisal, transfer, preservation and management of database, and to develop guidance advice and other supporting artefacts to support that approach."

Product 1: SIARD Case Study



Image Source: https://database-preservation.com/#desktop

Valuation Office Information System

- 2014 Closed agency transfer
 - Data file in native SQL (.MDF)
 - Log File in native SQL (.LDF)
 - Table data individually exported in XML
 - Technical documentation (Data Dictionary)
 - Screenshots of UI
- 2021 SIARD file



NATIONAL ARCHIVES OF AUSTRALIA

Checklist-of-questions-to-ask-an-Agency-with-data-to-transfer-froma-Relational-Database

4

R305582021¶

Digital-Archives-Research and Innovation 2021

4. Using this checklist¶

The-checklist-is-intended-to-assist-National-Archives'-staff-to-guide-discussions-with-agencies-to-come-to-a-shared-understanding-of-the-selection-and-transfer-process,-i.e.-appraisal,-pre-ingest-and-ingest-and-will-not-be-comprehensive-for-every-system-or-every-transfer.--¶

 $While \ the \ check list \ is \ detailed, \ it \ is \ important \ to \ note \ there \ are \ four \ main \ questions \ that \ will \ shape the \ form \ the \ transfer \ takes \ and \ which \ will \ influence \ which \ questions \ to \ focus \ the \ conversation \ on, \ and \ which \ will \ shape \ the \ transfer \ . \ \P$

- 1.→ Is-the-entire-database-RNA-or-only-part-of-it?¶
- 2.→ If-only-part-is-RNA, does-the-system-have-the-capacity-to-export-that-content?¶
- 3.→ If-a-partial-export-is-not-possible-or-it-is-all-RNA, can-it-export-the-entire-dataset?¶
- 4.→ Establish·if-periodic·snapshots·are·required,·i.e.·does·the·system·retain·all·changes·and·have·roll-back·capability,·or·is·the·content·subject·to·ongoing·change?··If·subject·to·change,·eg·data·is·overwritten·or·permanently·deleted,·periodic·transfer·will-be·necessary.¶

The investigation may involve several meetings involving different staff and roles, especially for more complex databases/business systems. Initial meetings contact usually take place with the agency IM manager/contact. Later meetings will include IT and will discuss system functionality/capability in detail.

 $It \emph{-} is \emph{-} recommended \cdot that \cdot the \emph{-} information \cdot gathered \cdot in \emph{-} response \cdot to \cdot these \cdot questions \cdot is \cdot collaboratively documented \cdot and \cdot shared \cdot between \cdot National \cdot Archives \cdot and \cdot the \cdot agency, \cdot so \cdot it \cdot can \cdot be \cdot built \cdot on \cdot as \cdot the \cdot transfer \cdot progresses. \P$



NATIONAL ARCHIVES OF AUSTRALIA

DETERMINING-OPTIONS-FOR-TRANSFERRING-DATABASES: GUIDANCE-FOR-NATIONAL-ARCHIVES'-STAFF¶

Digital-Archives-Research-and-Innovation-2021¶

*Determining-frequency-of-transfer¶

This-section-lists-criteria-for-determining-whether-a-transfer-of-data-contained-in-a-database-is-a-single-one-off-transfer-or-a-periodic-snapshot-carried-out-on-an-agreed-frequency.

One-off-transfer¶

One-off-transfers-(or-end-of-life-transfers)-are-usually-made-in-the-following-situations:

- → Closed-agency¶
- ◆ Machinery-of-Government-change, for example the inheriting agency may decide to transfer the database to the National Archives
- → Decommissioning-of-system¶
- At·the·point·of·migration·to·a·new·system·or·platform¶
- → Expiry-of-software-or-system-licence-or-support¶

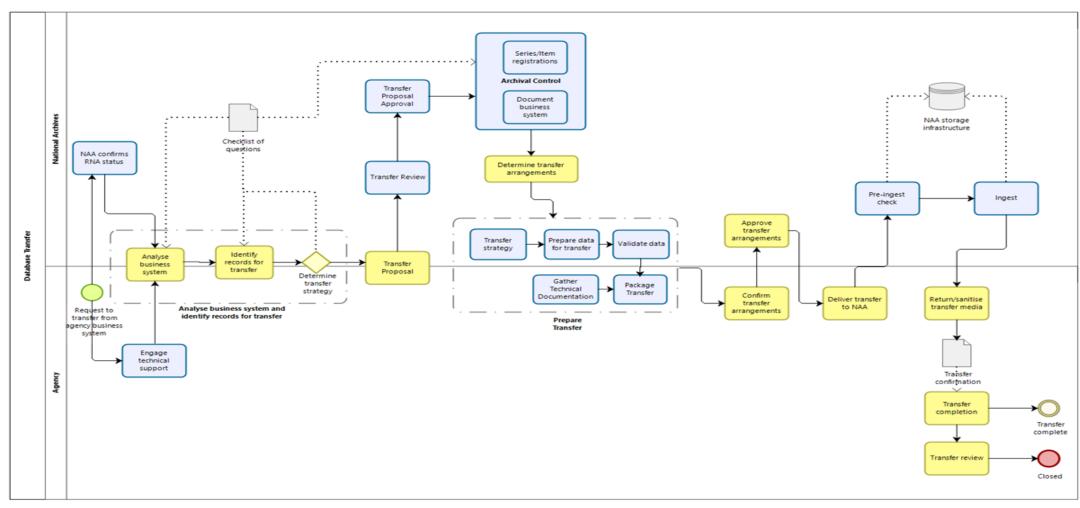
Periodic-snapshot¶

 $Periodic \cdot snapshots \cdot are \cdot made \cdot in \cdot the \cdot following \cdot situations \cdot and \cdot need \cdot to \cdot be \cdot appropriately \cdot planned \cdot and managed: \P$

- ◆ Business-system-periodically-deletes-or-overwrites-data-without-capturing-an-accurate-and-complete-log-of-data-changes¶
- → Business-system-is-highly-dynamic¶
- → Business-system-has-on-going-business-use, but-high-value-or-risk-has-led-to-a-decision-totransfer-periodically-to-the-national-Archives.¶

Attention-needs-to-be-given-to-frequency-of-transfer, and-whether-the-transfer-is-incremental-(i.e.-the-changes-between-snapshots)-or-a-complete-capture-of-content.--Frequency-of-transfer-would-be-determined-by-factors-such-as-when-data-is-deleted-or-overwritten-(e.g.-every-five-years),-or-when-there-is-significant-change,-for-example-an-annual-transfer-may-be-required-for-a-registration-database-that-is-updated-and-reviewed-at-the-end-of-the-financial-year.¶

Transfer Process Maps





Additional Metadata for Database/Business Systems

Purpose¶

This-document-identifies-additional-descriptive-metadata-elements-that-should-be-captured-to-better-describe-the-technical-structure-and-functionality-of-databases/business-systems.··While-the-CRS-system-adequately-captures-the-recordkeeping-context-of-record-creation-and-management,-digital-systems-require-at-least-some-of-their-technical-components-and-properties-to-be-identified-and-described-so-that-they-can-be-properly-understood-into-the-future,-long-after-they-have-ceased-to-exist.··This-metadata-is-also-important-when-providing-access-to-the-information-and-also-for-preservation-planning.··¶

 $Currently, these \cdot additional \cdot elements \cdot could \cdot be \cdot added \cdot to \cdot Series \cdot and / or \cdot Item \cdot Notes \cdot in \cdot RecordSearch, \cdot and \cdot could \cdot be \cdot captured \cdot in \cdot discussions \cdot with \cdot agencies \cdot when gathering \cdot information \cdot about \cdot the \cdot database / business \cdot system \cdot \cdot The \cdot Series \cdot Registration \cdot Form \cdot could \cdot be \cdot modified \cdot to \cdot include \cdot these \cdot elements \cdot \cdot In \cdot the \cdot full ness \cdot of time, \cdot as \cdot work \cdot continues \cdot on \cdot the \cdot National \cdot Archives' \cdot business \cdot system \cdot functionality \cdot can \cdot be \cdot improved \cdot to \cdot better \cdot accommodate \cdot these \cdot elements \cdot \P$

As-with-other-description-activity, the-professional-judgement-of-the-transfer-officer/description-officer/archivist-should-be-used-when-determining-what-descriptive-information-to-capture. This-document-is-provided-as-a-guide-and-prompt. ¶

 $The \cdot Standards \cdot Mapping \cdot column \cdot links \cdot to \cdot the \cdot same \cdot or \cdot similar \cdot properties \cdot in \cdot other \cdot relevant \cdot standards \cdot Archival \cdot Control \cdot Model \cdot (ACM), \cdot Data \cdot Dictionary \cdot for \cdot Preservation \cdot Metadata \cdot (PREMIS), \cdot Australian \cdot Government \cdot Recordkeeping \cdot Metadata \cdot Standard \cdot (AGRKMS), \cdot Software \cdot Metadata \cdot Recommended \cdot Format \cdot Guide \cdot (SMRFG). \P$

Metadata-Flements¶

Field¶	Description¤	Comment¤	Series-or-Item¤	Standards-mapping¤	¤
្ត Database-Management- Systemជ	What-database-management- system-including-version-and- date,-was-used,-e.gSQL- Server,-MySQL,-Oracle,- PostgreSQL-etc?¤	Important-for-access-and- contextImportant-if- emulation-is-an-access- option.¤	Series-Note¤	ACM:·21.3·Creating· Application·Name¶ ACM:·21.4·Creating· Application·Version¶ PREMIS:·1.5.5.1·Creating·	п

Project Outcome: Transferring Databases

- No one-size-fits-all approach
- Interpretation of the disposal class/analysis of business system
- Transfer could comprise a combination of:
 - SIARD file
 - Native SQL files
 - Export in an open data format like CSV, XML, etc.
 - Transfer of reports/views
 - Documenting how the business system was used
 - Technical documentation data dictionary, data model
 - Full descriptive information documented in the archival control system

Future Work

- Continue to develop knowledge and capability
- Pilot database transfers Microsoft Dymanics,
 Lotus Notes
- Work with agencies to imbed transfer decisions in the Records Authority creation process

Thank You