

CS 14 Shared Drive Migration

InterPARES 3
International Symposium – Oslo, Norway
2010 09 17

Lois Enns, Records Manager Legislative Services Division City of Surrey



Introduction

Repeatable process to analyze unstructured files and migrate records to an electronic content management system.

Two years to develop. Successful proof-of-concept: 42,000 files; 34,000 files classified as records, renamed, and migrated; 45 business days; 1 FTE plus tech support.

Estimate two years to analyze and migrate the balance: 27 shared drives; 2.9 million files; 2 million++ records.

This presentation is an overview of the Share Care process.



Shared Drive Migration Toolkit

The manual has seven sections; two sections apply to the organization:

- 1. Defining the organizational context
- 2. Defining the technical context

Five sections apply to each business unit:

- 3. Defining the business context
- 4. Completing a business analysis
- 5. Completing a technical analysis
- 6. Completing a classification analysis
- 7. Migrating records

Appendices:

Templates, Utility Applications Guide, Glossary, Bibliography



Corporate Records Manual





1.0 Defining the Organizational Context

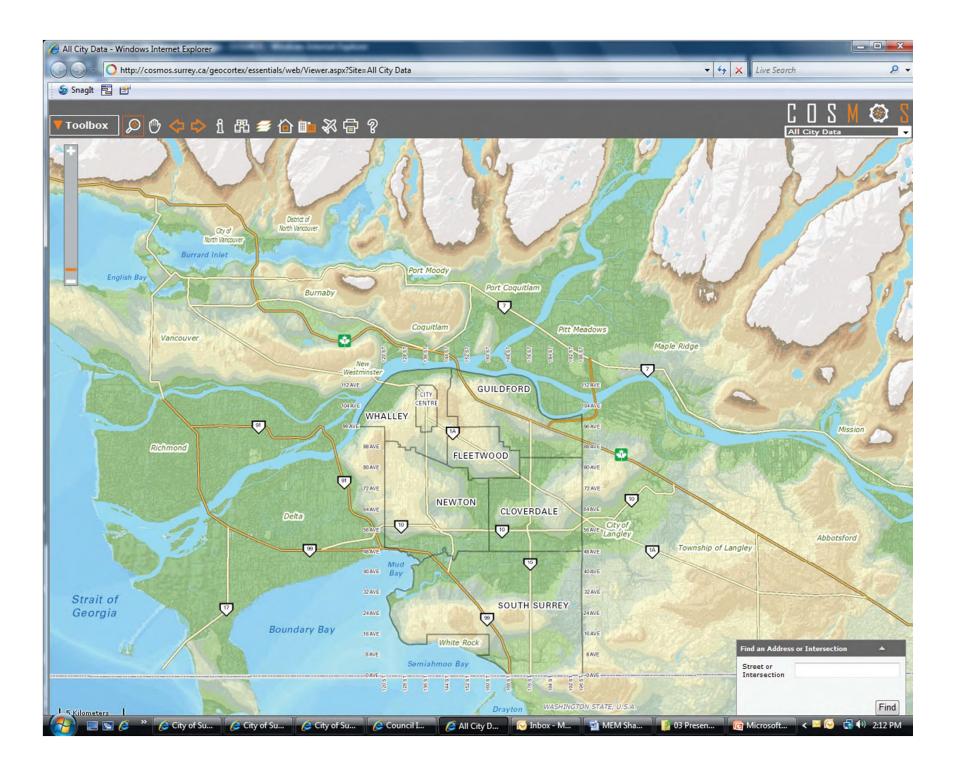


1.0 Defining the organizational context

- 1. Reviewing the legislative context
- 2. Reviewing the *Corporate Records By-law*
- 3. Reviewing staff responsibilities
- 4. Defining corporate records
- 5. Defining electronic records requirements
- 6. Reviewing records classification
- 7. Reviewing the records lifecycle
- 8. Defining transitory records

For this presentation, a more external view of the organization is provided.







History

- 1879 Incorporated as a municipality; logging and farming
- 1969 Sons of Norway's Viking Lodge No. 104 formed
- 1974 Agriculture Land Reserve established
- 1993 Becomes a city of 100,000+
- 2008 Population reaches 452,000
 - Second largest city in BC
 - Twelfth largest in Canada
 - Growing at 1,000 people a month
 - Mixed use: industrial, commercial, residential, farming
- 2010 Participates as an Olympic venue city
- 2025 Expected to become the largest city in the region



Organizational culture

- Low ratio of employees to citizens
 - CoS = 452,000 citizens : 3300 employees
 - CoV = 578,000 citizens : 9000 employees
- Stable environment
 - Same Mayor and Council for two terms (and longer)
 - Many long-term employees
 - Promotion from within
 - Cooperative union environment
- Tradition of a practical "journeyman" approach now overlaid with professionals
- "Lean and mean"
 - Citizens want value for dollars; employees are also citizens







Records environment

- 1966 First file plan adopted as policy
- 1979 Records Centre opens; microfilming begins
- 1992 Records retention by-law; first shared drives implemented
- 1996 Provincial Freedom of Information and Protection of Privacy Act
- 1999 MOA functional classification adopted; first manual published
- 2000s Large database systems implemented
- 2005 Electronic records overtake paper; scanning program begins
- 2008 RM program renewal; ECM selection process begins
- 2010 Corporate Records By-law adopted; first ECM implementation



"Everything all at once"

- Formalize the records program
 - Records Census (paper)
 - Updated Corporate Records By-law
 - Publish the Corporate Records Manual (four volumes)
 - Build staff capacity (Records Centre, Central File Areas)
 - Develop a corporate-wide training program (CBT)
- Contribute to the ECM project
 - Requirements, RFP, selection committee, Council Report, contract
 - Gap analysis, use case testing, design decisions, Model Office
 - Migration of all unstructured content stored on shared drives
 - Change management, user training, support
 - Maintenance activities (information model, metadata model, permissions management, electronic disposition...)



2.0 Defining the Technical Context



2. Defining the technical context

- 1. Conducting a server inventory
- 2. Analyzing server roles
- 3. Conducting a shared drive inventory
- 4. Identifying shared drives for migration



Scope definition

- 310 servers
 - 907 shared drives
 - 26+ shared drives of primary interest
 - 2.9 million files, 1.8 terabytes



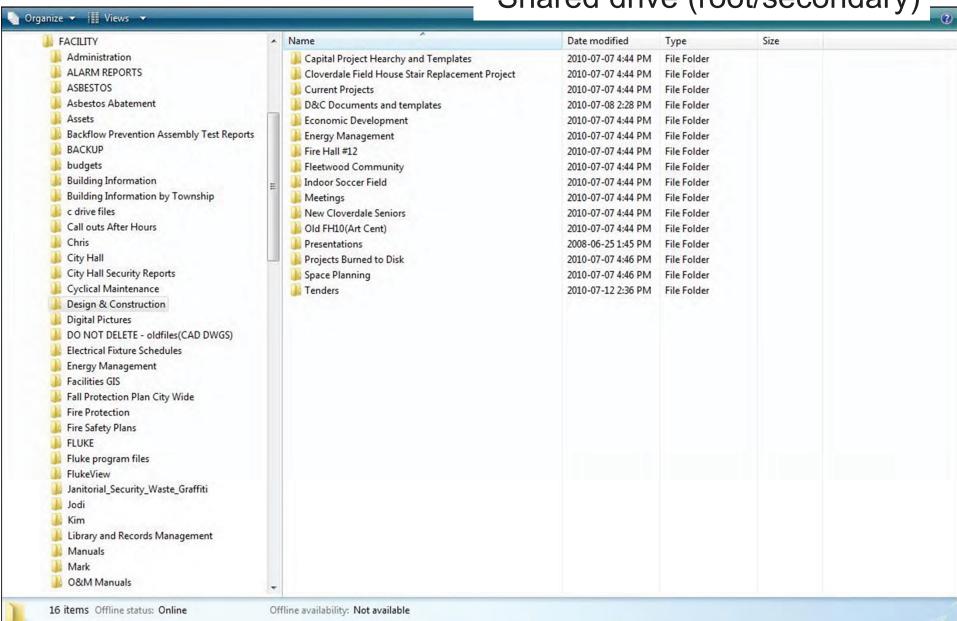
3.0 Defining the Business Context

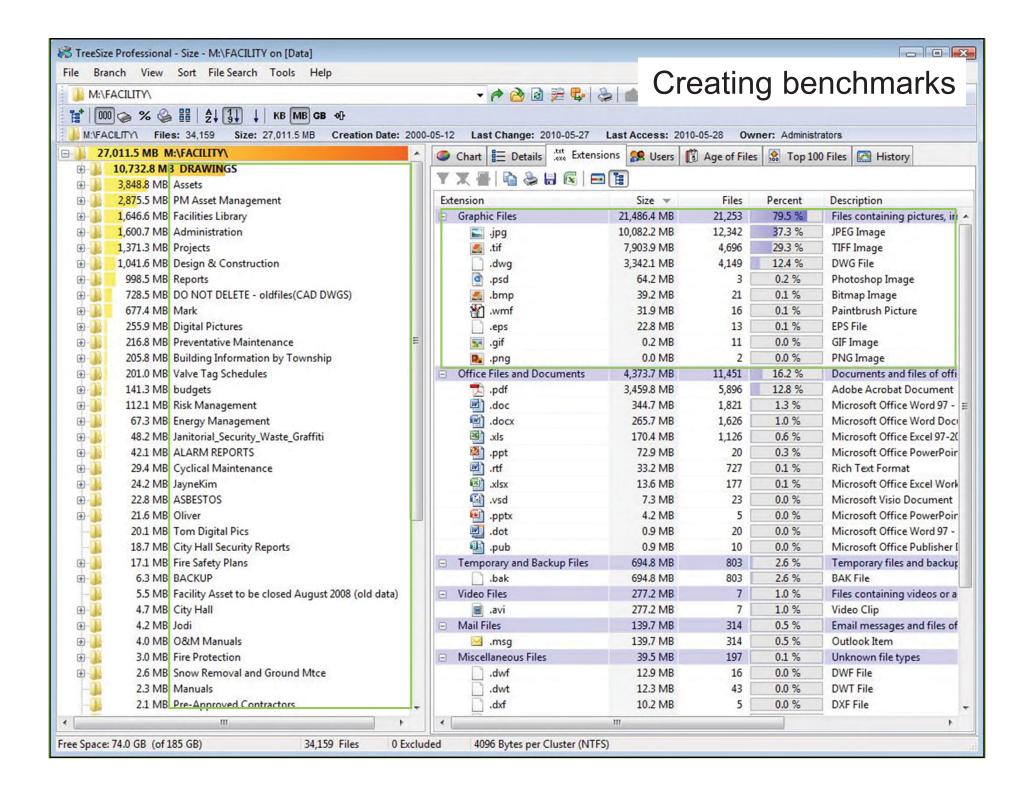


3 Defining the business context

- 1. Creating an administrative history
- 2. Creating a record history
- 3. Creating shared drive benchmarks
- 4. Creating a user history
- 5. Creating private storage benchmarks
- 6. Updating the Corporate File Plan

Shared drive (root/secondary)





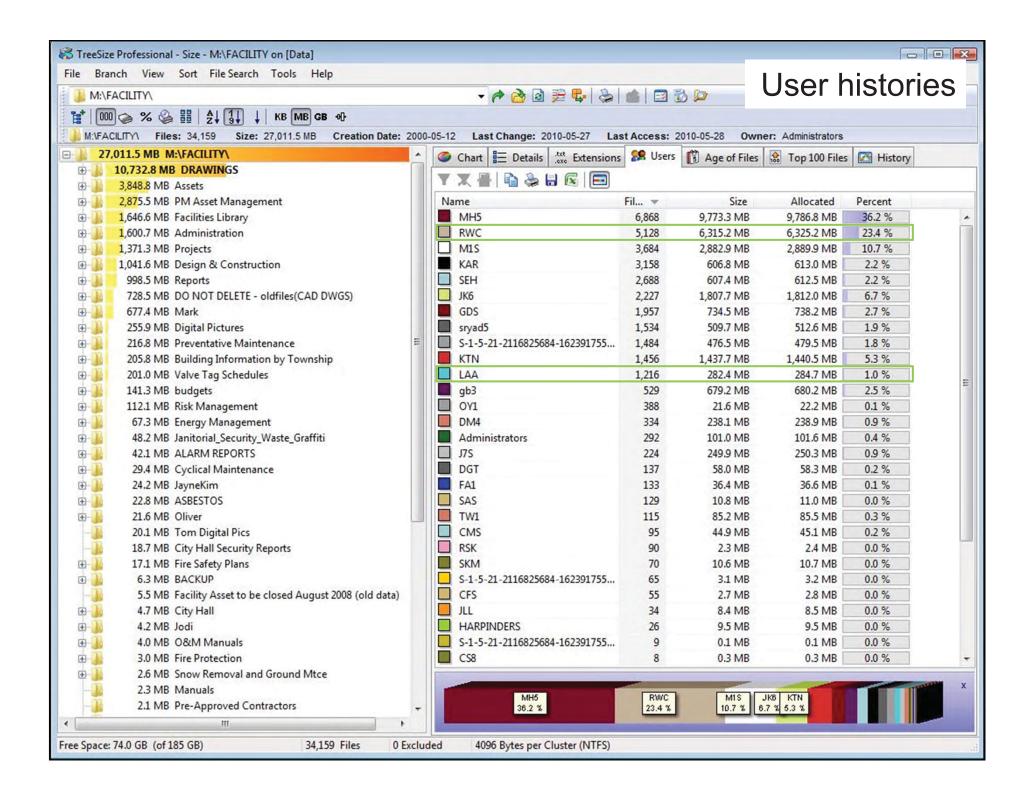
Shared Drive Summary

FOLDER OWNERS

FOLDER NAME	SUBFOLDERS	FILES	DISI
Administration	227	3219	1.70
Alarm Reports	191	1879	0.10
Alex Scott	0	29	3.6 N
Asbestos	18	55	29.1
Asbestos Abatement	0	3	0.1 N
Assets	551	3939	4.5 0
Backflows Prevention	0	3	0.2 N
Backup	6	54	50.
Budgets	165	845	0.2 0
Building Information	182	175	0.2
Call Outs After Hours	2	21	1.9 N
Chris	2	27	0.6 N
City Hall	2	7	4.7 N
City Hall Security Reports	0	61	18 N
Cyclical Maintenance	7.	73	29 N
Design and Contruction	193	3847	1.60
Diane George	3	10	4.7 N
Digital Pictures	1	31	0.20
Do Not Delete (old files)	227	4613	1.50
Drawings	1227	9666	10.7
Electrical Fixture Schedules	9	5	0.1 N
Energy Management	34	176	0.1
Facilities Library	283	3364	2.20

Shared Drive Format Summary

YEAR	JPG	PDF	DWG	TIF	DOC(X)	XLS(X)	RTF	ВМР
2010	734	482	7	2	159	81	0	0
2009	4,154	1,831	46	6	1,774	222	2	0
2008	3,862	1,308	344	8	209	335	21	2
2007	1,108	1,564	891	139	440	296	695	1
2006	317	297	905	253	141	38	3	4
2005	1,259	243	159	163	77	32	0	1
2004	541	127	62	189	203	60	0	0
2003	294	36	124	3,546	100	49	1	3
2002	56	7	789	389	100	40	3	7
2001	0	0	69	0	173	72	0	0
2000	0	1	324	0	65	33	2	1
1999	0	0	134	1	3	43	0	1
1998	0	0	97	0	0	0	0	1
1997	0	0	126	0	2	1	0	0
1996	0	0	72	0	0	1	0	0
1995	0	0	0	0	1	0	0	0
1994	0	0	0	0	0	0	0	0
1993	0	0	0	0	0	0	0	0
1992 & Prior	17	0	0	0	0	0	0	0
Total	12,342	5,896	4,149	4,696	3,447	1,303	727	21
Disk	9.8GB	3.4GB	3.3GB	7.7 GB	0.6GB	0.2GB	0.05GB	0.00G



User History

SHARED DRIVE NAME: FACILITIES OFFICE OF PRIMARY RESPONSIBILITY: CIVIC FACILITIES DIVISION

ID	ACTIVE	USER NAME	JOB TITLE	NO. FILES	USER GROUP	DPT	DIV	SCT	LOC	PREMISSIONS
Реверия Реверия	YES		System Admin	765						
Axxx9	YES	Bob Martin	System Admin	1828		FIT	ITS	ISR		
Axax9	YES	lain Malcolm	System Admin	4		FIT	ITS	ISR		
Аххаа9	YES	Seaton Mahanger	System Admin	1		FIT	ITS	ISR		
Axxa9	NO	Orphaned user		2098						
Axaaa9	NO	Orphaned user		66	-					
Axxax9	NO	Orphaned user		9						
Axaxa9	NO	Orphaned user		7						
Axxxx9	NO	Orphaned user		3	0,-					
Аааах9	NO	Orphaned user		1						
Aaaxx9	NO	Orphaned user		1						
ARCADA Admin	NO	(Service Account)		7						(Backups)
RWC	YES	Bill Tan	Engineering Assistant	6,783		PDV	CFC			R+W+D
M1S	YES	Lois Murphy		3,921		ENG				R+W+D
KAR	YES	Jill Krop	Clerk	3,369		PDV	CFC			R+W+D

Page 3-20 | City of Surrey

FACILITY AND PROPERTY MANAGEMENT

					0700 – 0999	
		DISCO	0710		Facility Administration	
0712	0	Buildings, Facilities and P – Emergency Drills	0720	Ż	Cyclical Replacement and Renovations	
0715	0	Accommodation – Gener	0730		Facility Service Agreements	
0770	0	Buildings – Damages	0740	È	Facility Security	Limited Access (Restricted)
0780	0	Buildings – Disposal	0750		Preventative Maintenance	
0790	0	Buildings – Maintenance	0755	Ė	Facility Assets	
0800	0	Buildings – Individual	0760		Design and Construction	Limited Access (Confidential)
0820	0	Facilities – General	0775	ź	Facility Building Plans	Limited Access (Confidential)
12000000	_		0785	ź	Fire Safety Plans	
0890	0	City Properties – Mainter	0850		City Property Administration	
0970	0	Utilities – General	0870		City Property Acquisitions	
0975	0	Residential Electricity Inf	0880		City Property Development	
		1	0910		City Property Sales	9 series – no change
			0930		Leased Properties	5 series – new 10 series – cancelled
					L. Carrier L.	

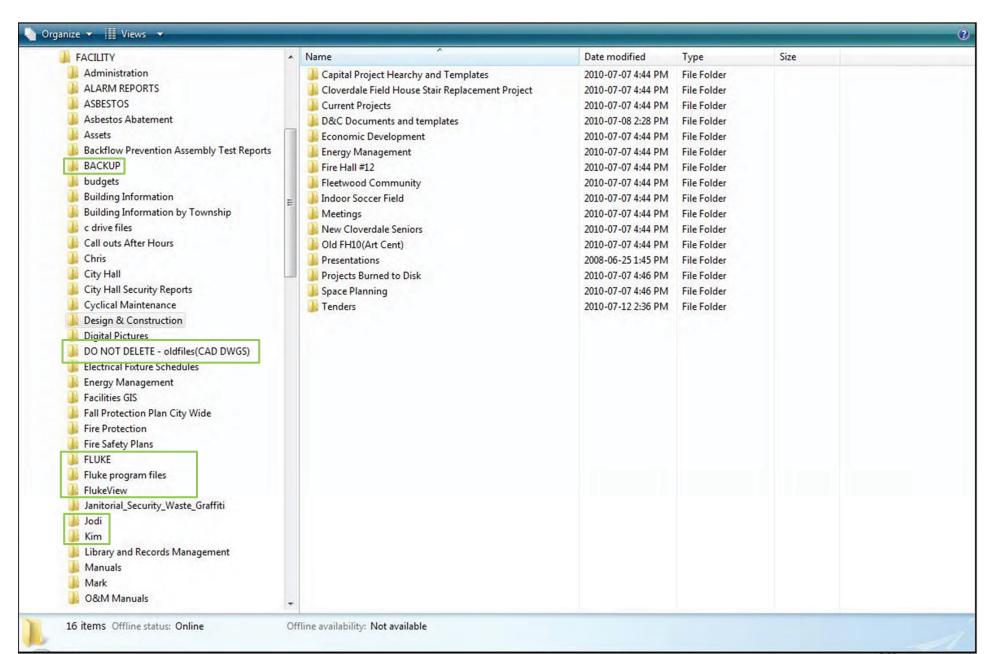


4.0 Completing a Business Analysis



4.0 Completing a business analysis

- 1. Providing a change management framework
- 2. Launching a Share Care project
- 3. Reviewing private storage
- 4. Reviewing shared storage



Root Folder Summary

Folder Name	No. Files	Description	Decision
C Drive Files	541	Unused system and application files	Move to NON-BUSINESS FILES folder
Do Not Delete – Old Files (CAD DWGS)	3,484	Engineering Assistant notes that the folder results from an earlier project designed to organize the Drawing folder. Suggests that the files be reviewed at a later date to determine if the folder contains any final or as-built copies. Believes the folder likely contains preliminary and working drawings.	After renaming and classifying files in the Drawing root folder, review to determine any missing sets. Add any missing sets to FACILITY BUILDING PLANS. Move any duplicate sets to DUPLICATES folder.
		Further review revealed that sub- folders Fire Halls, Police, City Hall, Cloverdale, Fleetwood, Guildford, Newton containing about 1600 files were organized and the South Surrey and Whalley folders were partially completed.	
Fluke	178	Engineering Assistant notes that the Fluke program was formerly kept on the shared drive to allow staff to	Move to NON-BUSINESS FILES FOLDER
Fluke Program Files	6	upload the program to their own drives.	Move to NON-BUSINESS FILES FOLDER
Elukeview	6	Under current policy, staff are not allowed to load their own programs and must send a request to ITS Help Desk instead.	Move to NON-BUSINESS FILES FOLDER
Asset Management/ Asset Inventory	3,135	Engineering Assistant notes that individual inventory sheets (docs) are no longer required as they have been summarized in the Asset Master List (XIS).	Move Asset Sheets to OBSOLETE RECORDS folder (300 files). Rename and classify the Images (2800 files) and other files as appropriate.
		Manager notes that the related photos will eventually reside in the City Works database environment.	
Facilities Library/ Oscars Files	1,823	Folder was created by a staff person no longer with the City. Manager notes that the staff person	Rename and classify to FACILITY BUILDING PLANS.
		provided digital copies of as-builts to other staff members.	



5.0 Completing a Technical Analysis

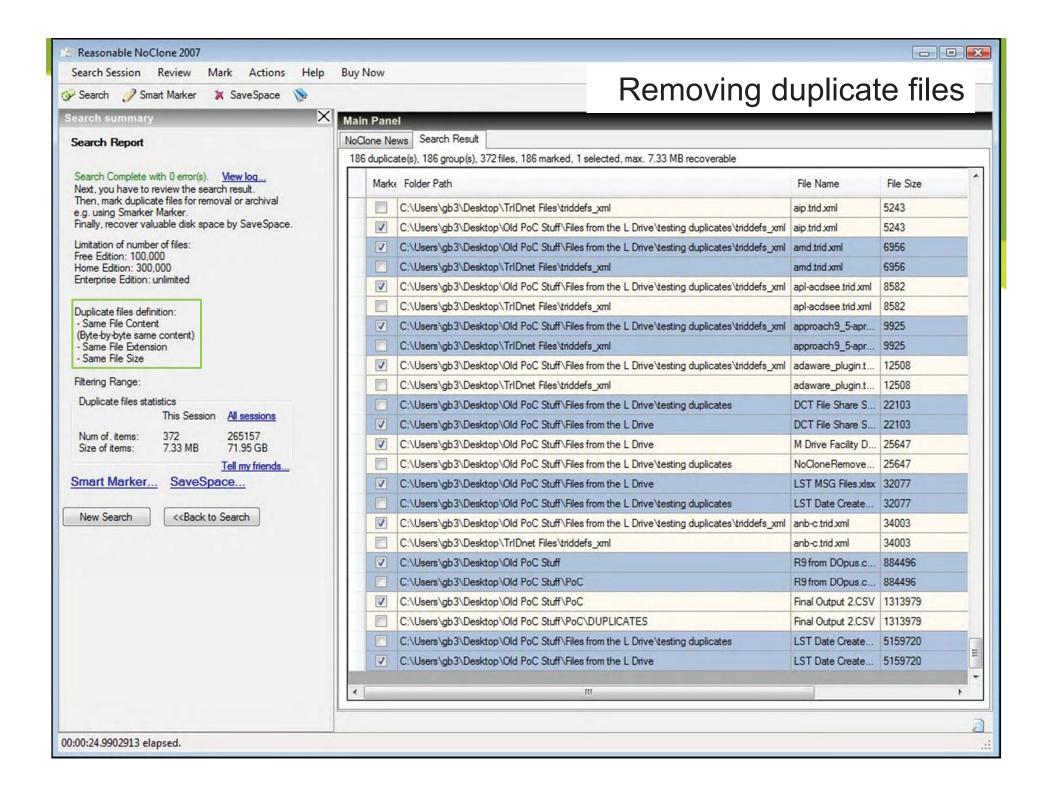


5.0 Completing a technical analysis

- 1. Preparing an appraisal environment
- 2. Removing business discards
- 3. Extracting zip files
- 4. Resolving long file names
- 5. Removing duplicate files
- 6. Evaluating file formats
- 7. Removing empty folders

Shared Drive Tracking Sheet

Sub-procedure	SDMT section	Restore	Discards	Files	Total Files	Total Size (MB)
Appraisal Copy	5.1	R0	None	42,361	42,361	32,068.2
Designated duplicate files			- 1	-1,479		-488.4
Non-business files				-1,053		-751.1
Obsolete records				-380		-682.5
Business Discards Removed	4.4, 5.2	R1	D1	42,361 -4,212 =	39,449	30,146.2
Zip files				-93		-443.4
Zip-within-zip files				-2		-4.6
Extracted files				1044		766.1
Zip Discards Removed	5.3	R2	D2	39,449 - (93+2) + 1,044 =	40,398	30,464.3
Long File Names	5.4	R3	None	56	40,398	30,464.3
Duplicate files (originals and duplicates)				5,996		
Originals kept				2,710		-
Duplicates removed				-3,286		-1213
Duplicates Removed	5.5	R4	D4	40,398 - 3,286 =	37,112	29,251.3
Files with missing extensions			1	59	59	
Files renamed				58	58	
Files with unidentified formats				1	1	
Files to be removed				0	0	
[Missing Extensions Resolved]	5.6	(R5)	(D5)	37,112	37,112	29,251.3

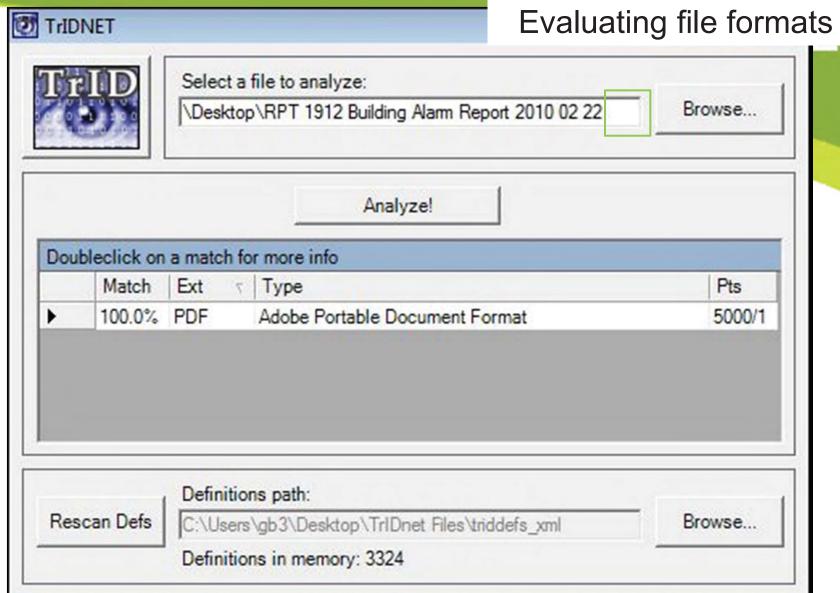


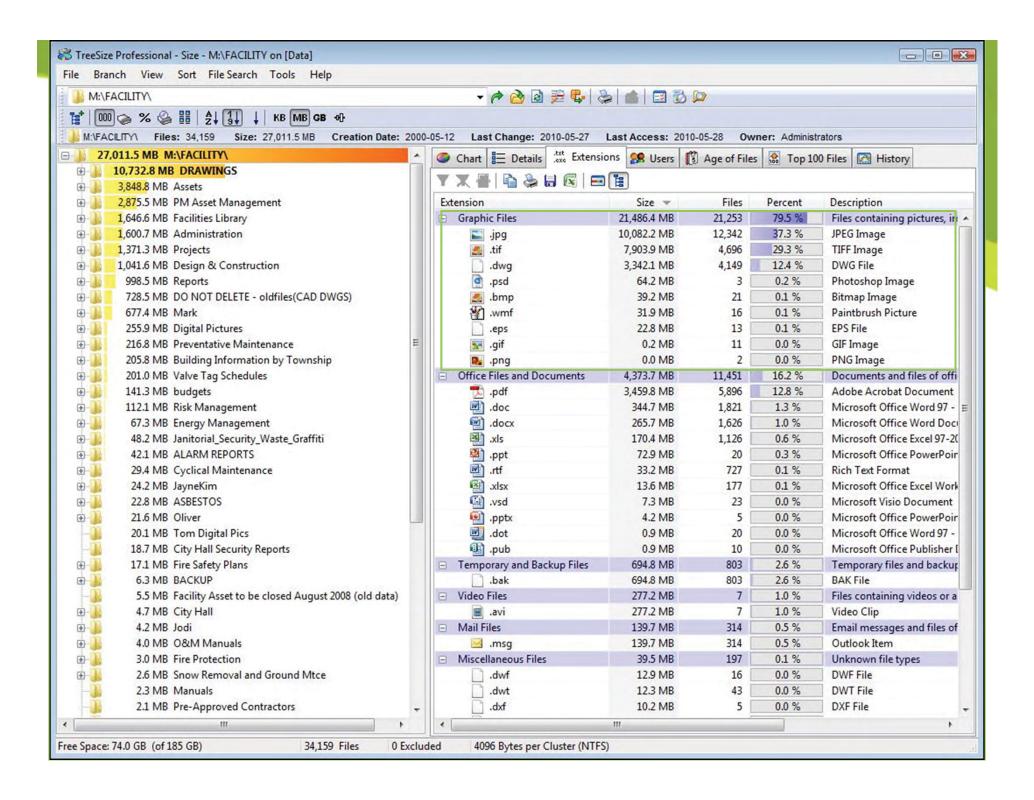


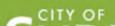
Shared Drive Tracking Sheet

Sub-procedure	SDMT section	Restore	Discards	Files	Total Files	Total Size (MB)
Appraisal Copy	5.1	R0	None	42,361	42,361	32,068.2
Designated duplicate files				-1,479		-488.4
Non-business files				-1,053		-751.1
Obsolete records				-380		-682.5
Business Discards Removed	4.4, 5.2	R1	D1	42,361 - 4,212 =	39,449	30,146.2
Zipfiles				-93		-443.4
Zip-within-zip files				-2		-4.6
Extracted files				1044		766.1
Zip Discards Removed	5.3	R2	D2	39,449 - (93+2) + 1,044 =	40,398	30,464.3
Long File Names	5.4	R3	None	56	40,398	30,464.3
Duplicate files (originals and duplicates)				5,996		
Originals kept				2,710		
Duplicates removed				-3,286		-1213
Duplicates Removed	5.5	R4	D4	40,398 - 3,286 =	37,112	29,251.3
Files with missing extensions				59	59	
Files renamed			F	58	58	
Files with unidentified formats				1	1	
Files to be removed				0	0	
[Missing Extensions Resolved]	5.6	(R5)	(D5)	37,112	37,112	29,251.3

CITY OF







FORMAT SUMMARY (After Appraisal)

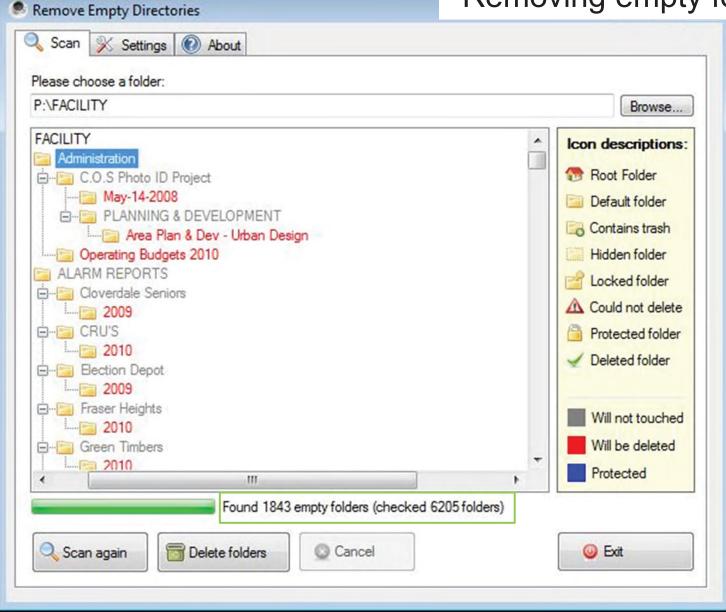
DECISION	DOCUMENT		IMAGE		AUDIO VISUAL	PROGRAM	PROGRAM SUPPORTING		DATABASE	COLLECTION	SYSTEM	INTERNET	TEMP/ BACKUP	EMAIL	MISC																			
MIGRATE (32)	DOC DOCX DOT MPP PDF PPT PPTX PUB RTF	TXT VSD XLS XLSM XLSM XLSX XLW	BMP CDC DGN DWF DWG DWT DXF EPS GIF JPG	PNG PSD PTN TIF WMF	AVI	AVI	AVI	AVI																VI									MSG	
STAY-IN- PLACE (26)						EXE	ADP ALB API DST FVF IMR LYR	PPR PRJ PST QRN SBN SBX SHP TFW		DBF DBI LDB MDB		AUX DAT DBG INI PC3		BAK																				
DESTROY (81)	ASC EXC LOG PM THMX WK1 WK3		CDR CMX CPT MXD PLT RED		ASX MOV	EX_ DLL	ALI AVE CGX COF CR5 CTB DWL DWL2 FMP FOT HTF ID INP	LK# LSW LOC LY MPD NFL ODB PCL PDT RF SHS	SLP TAB TBL	BDB CAT DB SHX TKB	PKG Z	CFG DMP DS_STOR E INF INS ISS LIB OPT TTF XML SYS	нтм нтт	AC\$ DFM DSC ERR FBK LNK TMP XLG XLK	BOX PRF XNK	386 CF DS\$ INFO JOB																		



Sub-procedure	SDMT section	Restore	Discards	Files	Total Files	Total Size (MB)
Files identified as stay-in-place				-358	-	-1633.4
[Stay-In-Place Files Resolved]	5.6	(R5)	(D5)	37,112 - 358 =	36,754	27,617.9
Formats identified for destruction				-2,596		-606.4
File Formats Removed	5.6	R5	D5	36,754 - 2,596 =	34,158	27,011.5
Folders identified as empty		-		-1,482 Folders		
Empty Folders Removed	5.7	R6	D6	5,893 - 1,482 = 4,411	34,158	27,011.5
Date created greater than date modified				20,273		
Author metadata present				6,640		
Company name present				4,035		
Author and Company Name Removed	6.1	R7	D7	0	34,158	27,011.5
Unique Identifier Applied	6.2	R8	None	All	34,158	27,011.5
Global Settings Applied	6.4	R9	None	All	34,158	27,011.5
Additional discards identified				-112		
Classify and Rename Complete	6.5	R10	D10	34,158-111	34,046	26,991.9

CITY OF

Removing empty folders





Sub-procedure	SDMT section	Restore	Discards	Files	Total Files	Total Size (MB)
Files identified as stay-in-place				-358	-	-1633.4
[Stay-In-Place Files Resolved]	5.6	(R5)	(D5)	37,112 - 358 =	36,754	27,617.9
Formats identified for destruction				-2,596		-606.4
File Formats Removed	5.6	R5	D5	36,754 - 2,596 =	34,158	27,011.5
Folders identified as empty				-1,482 Folders		
Empty Folders Removed	5.7	R6	D6	5,893 - 1,482 = 4,411	34,158	27,011.5
Date created greater than date modified				20,273		
Author metadata present				6,640		
Company name present				4,035		
Author and Company Name Removed	6.1	R7	D7	0	34,158	27,011.5
Unique Identifier Applied	6.2	R8	None	All	34,158	27,011.5
Global Settings Applied	6.4	R9	None	All	34,158	27,011.5
Additional discards identified			I	-112		
Classify and Rename Complete	6.5	R10	D10	34,158-111	34,046	26,991.9



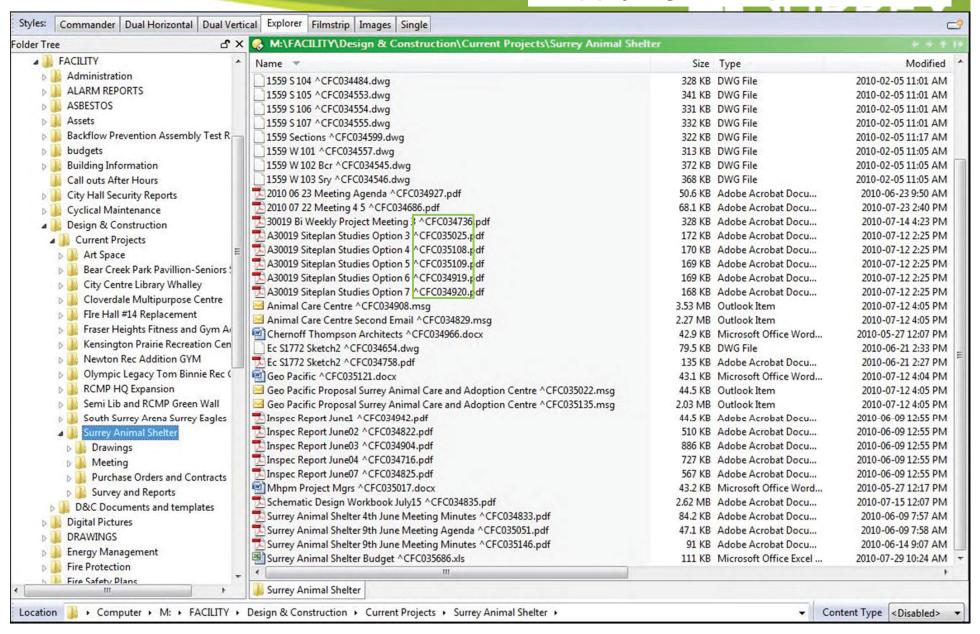
6.0 Completing a Classification Analysis



6.0 Completing a classification analysis

- 1. Reviewing file property metadata
- 2. Applying a unique identifier
- 3. Validating the naming standard
- 4. Applying global changes
- 5. Renaming and classifying records
- 6. Conducting quality assurance on named and classified records

Applying a unique identifier

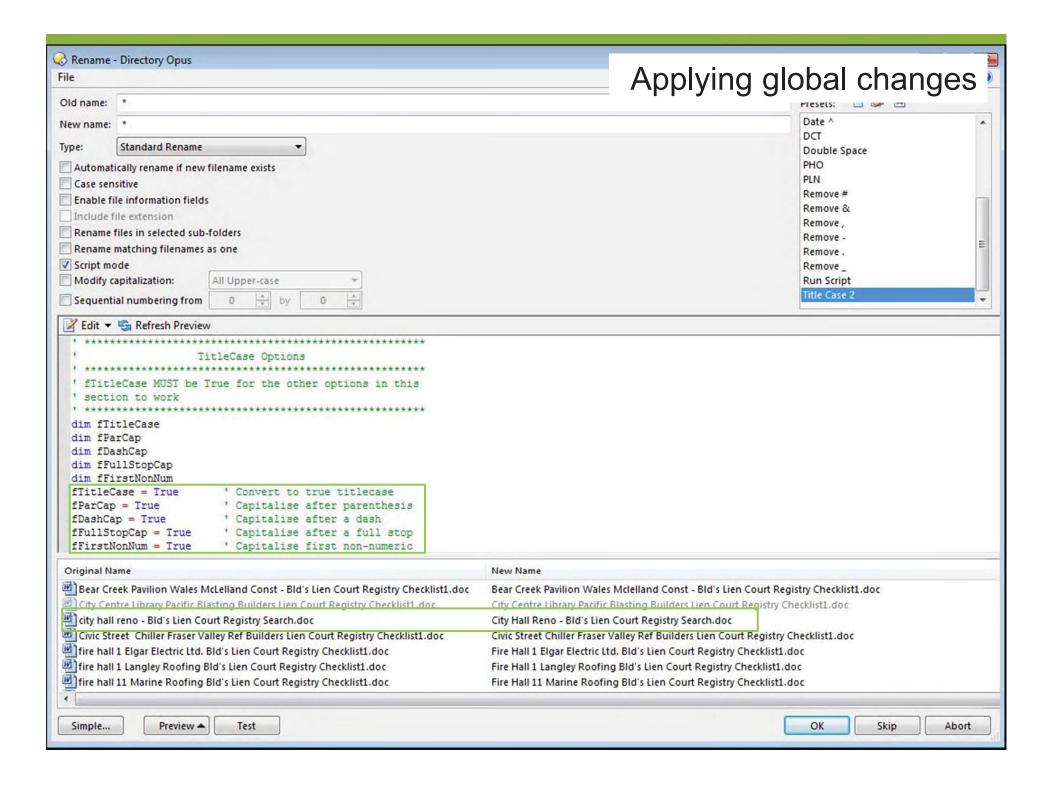


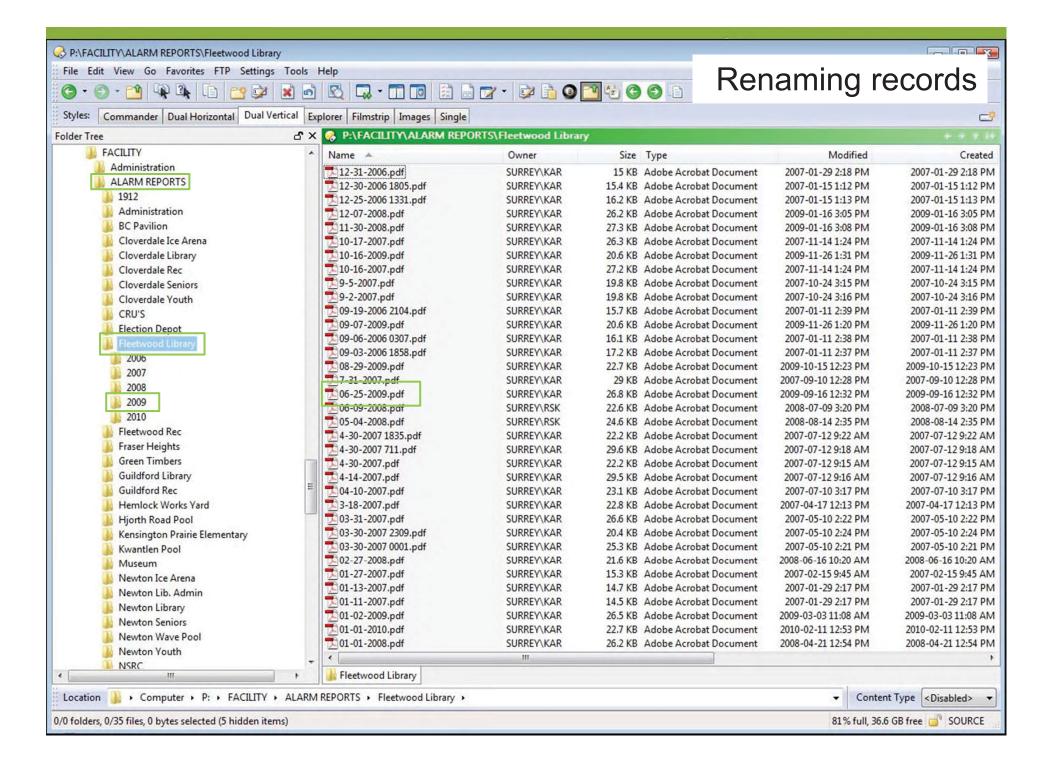


6.4 Implementing a record naming standard

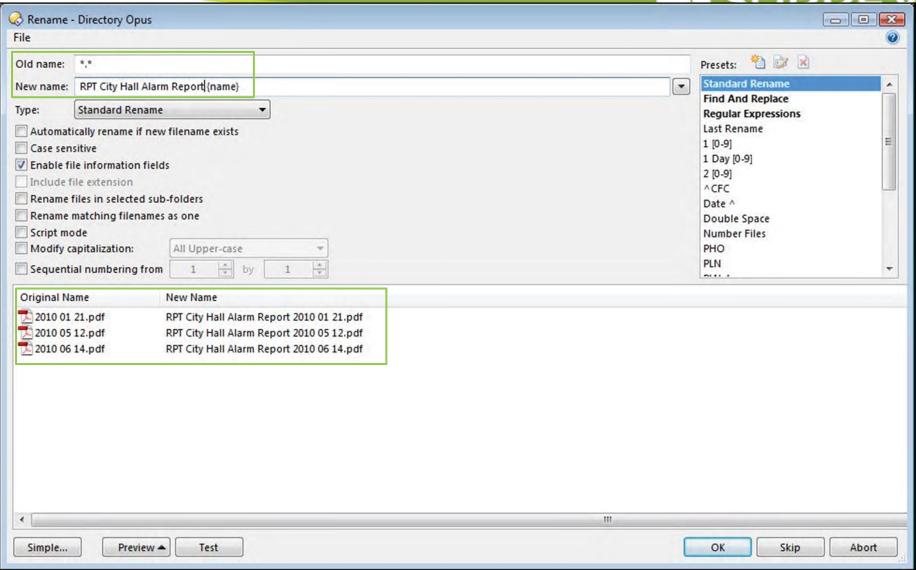
PLN-A 2009-1001 Proper Noun sA01 Subject Author YYYY MM DD v01-9.fmt

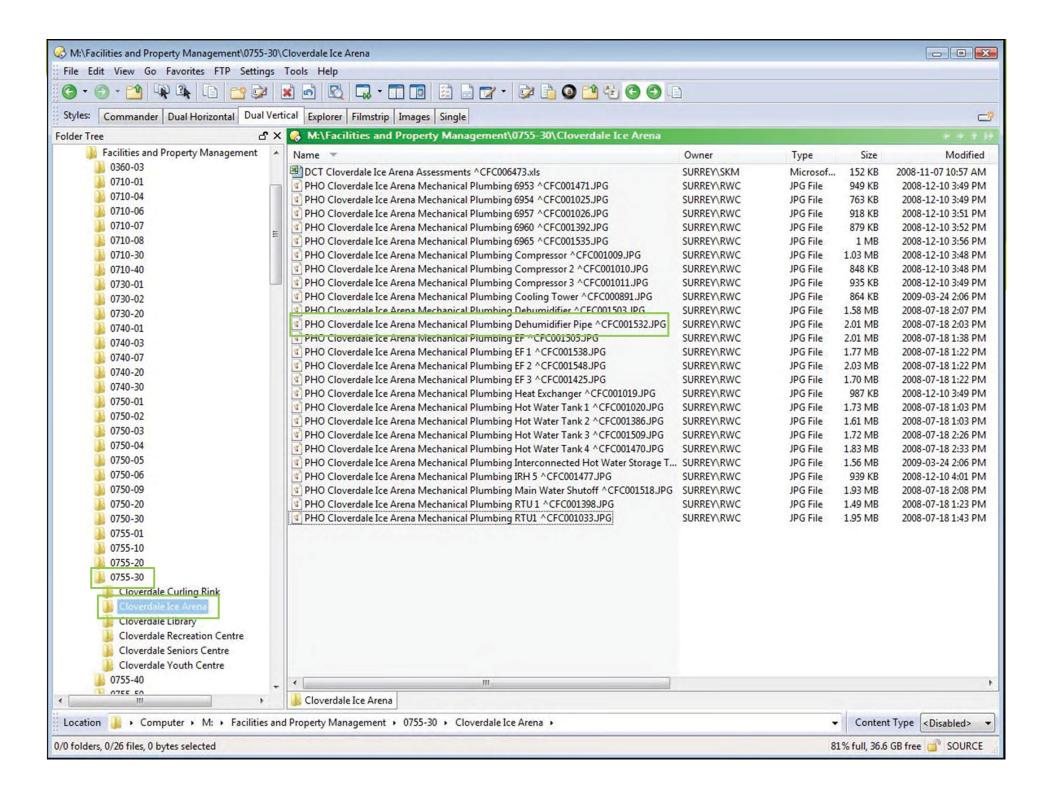
- 1. Record Type
- 2. Project Number
- 3. Proper Noun
- 4. Sequence Number
- 5. Subject
- 6. Author
- 7. Event Date
- 8. Version Number
- 9. <u>File Format</u>











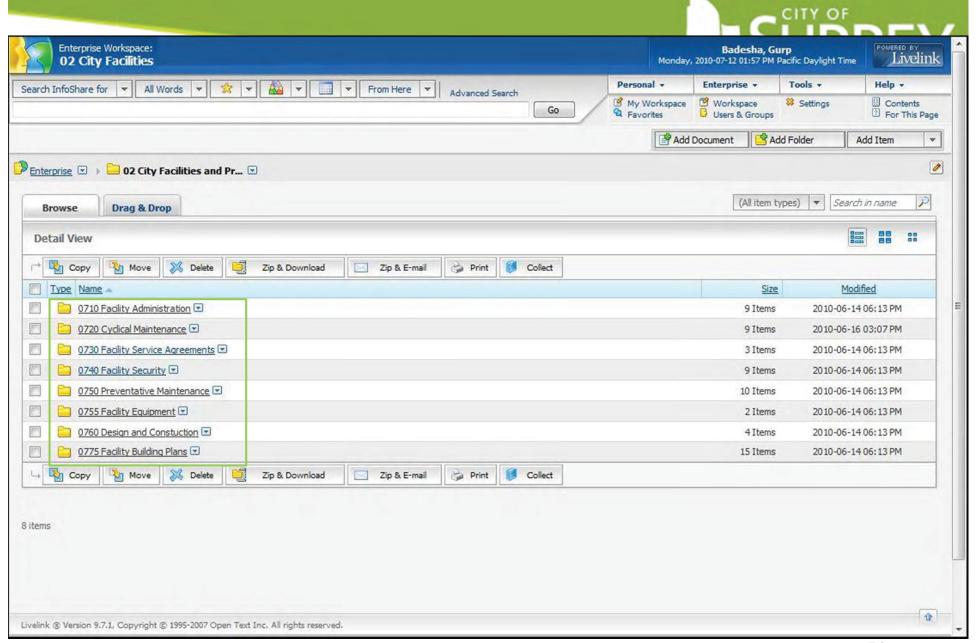


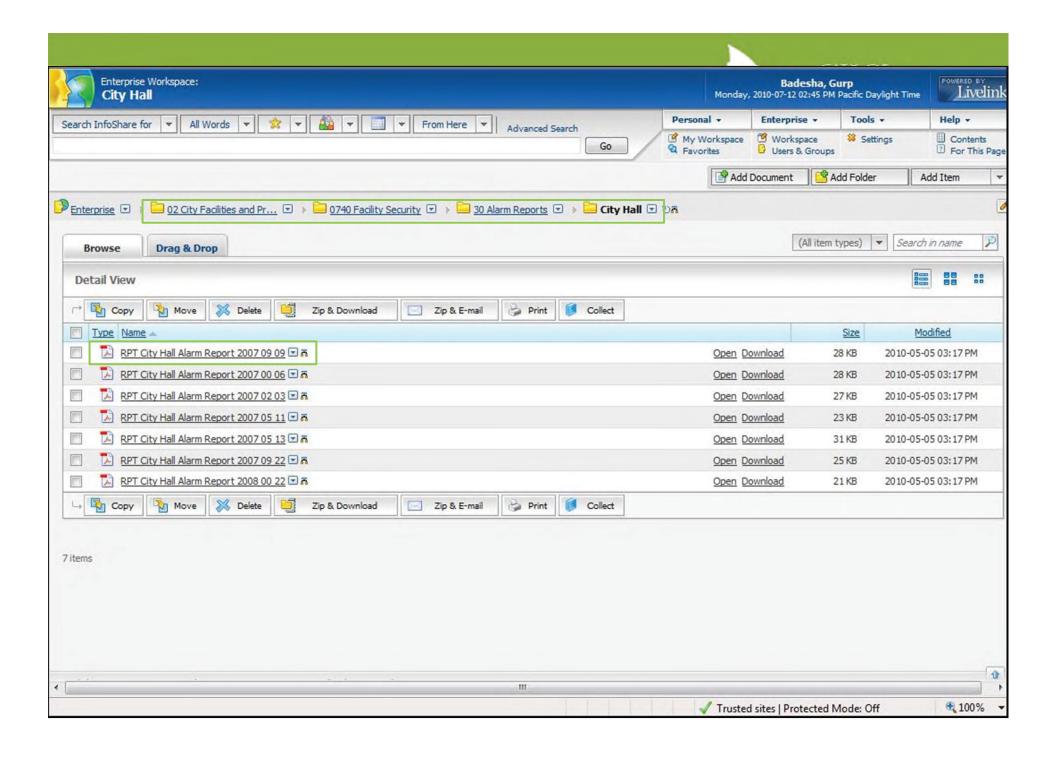
7.0 Migrating Records

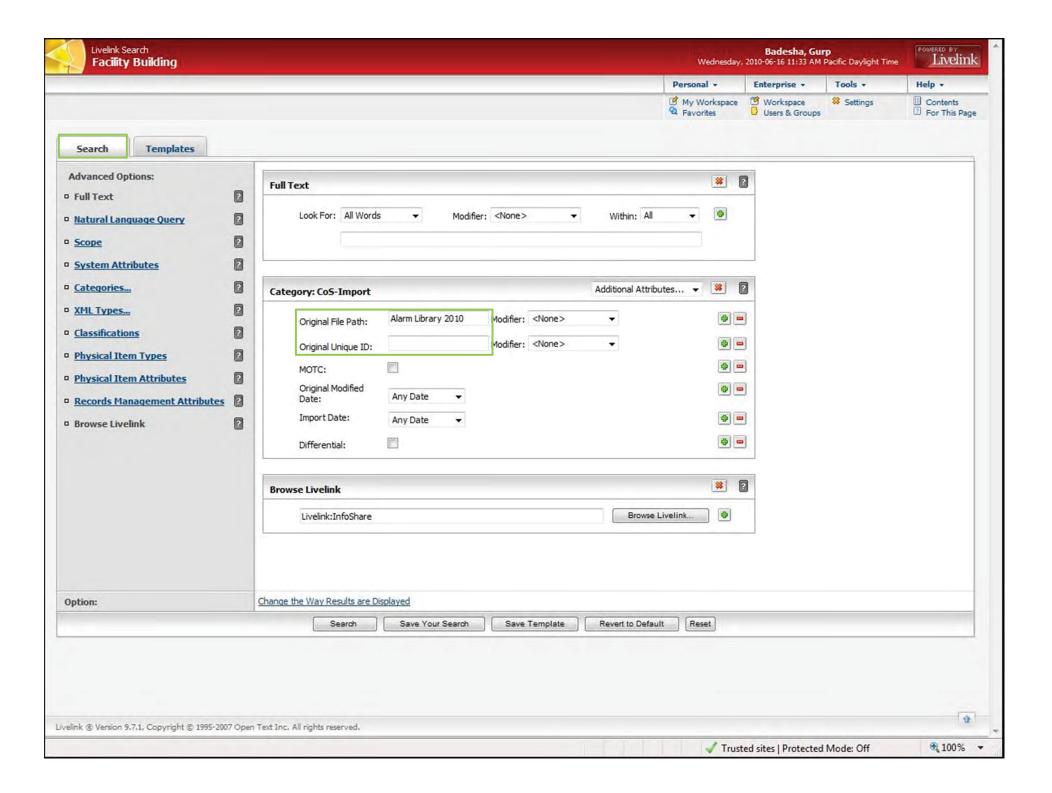


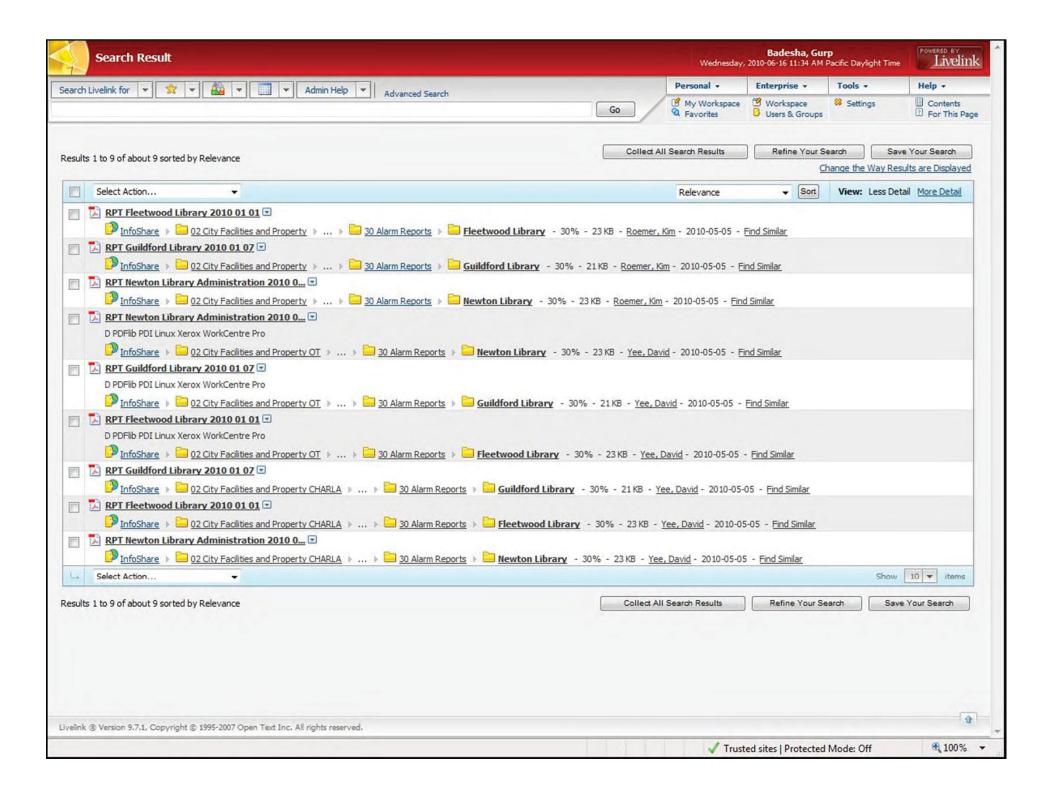
7.0 Migrating records

- 1. Migrating the appraisal copy
- 2. Conducting quality assurance on appraisal migration
- 3. Migrating the differential(s)
- 4. Decommissioning the shared drive
- 5. Migrating missing records
- 6. Destroying the shared drive backups











Closing Comments



Closing comments

- Think about whether or not the SDMT process applies
- Look at the business case, know how to sell the effort
- Recognize the technical and human resource requirements
- Expect that existing politics and trigger points will intensify



End