National Records of Scotland

Depositor Guidance for the Transfer of Archival Born Digital Records

May 2020

This Document was

Prepared by:	Eve Wright	Digital Records Archivist	5 May 2020
Reviewed by:	Garth Stewart	Head of Digital Records Unit	13 May 2020
Reviewed by:	Susan Corrigall	Head of Preservation	18 May 2020
Authorised by:	Laura Mitchell	Deputy Keeper of the Records of Scotland	27 May 2020

Document Distribution

Depositors of born digital archival records to NRS.
To be published on NRS website.

Amendment Suggestion

If you have any suggested amendments, please contact digital_records@nrscotland.gov.uk.

Status Control

Version	Date	Status	Prepared by	Reason for Amendment
1.1	5 May 2020	Final	Eve Wright	New version required due to changing requirements from depositors.

Contents

1 Introduction

This guidance document relates to the transfer of **born-digital records selected for permanent preservation** to National Records of Scotland (NRS).

Born digital records provide their own unique challenges in terms of transfer and preservation. In combination with support from your NRS Client Manager, this document will outline what record creators and NRS need to do to ensure born digital records can be transferred to, and preserved by, NRS.

2 Digital Preservation at NRS

NRS has been accepting born digital records since 1998. Our digital repository, overseen by the NRS Digital Records Unit (DRU), allows archival records to be preserved in an environment where preservation and security are paramount. NRS provides the technological expertise required to ensure long term preservation of records, administering checks and actions that may not be in place in record creators' own IT systems. These steps are essential to preserve and protect the reliability, integrity, authenticity and usability of these archival records, so that they can be accessed by future generations.

3 First steps

As part of the process of transferring records to NRS for permanent preservation, you will need to work closely with your assigned NRS Client Manager. Where the transfer involves any born digital material, whether it is solely digital or a hybrid between analogue and digital records, these conversations will be in collaboration with the DRU. As a general rule of thumb, the earlier that these discussions start the easier the transfer process is for all parties. We will provide technical guidance and support for the entire process of transfer. We will also need to gather some basic information from depositors in order to prepare the records for archiving and preservation and to deal with any technological challenges that may arise.

3.1 Your Recordkeeping System

One of the first questions we ask our depositors is what system is currently being used to store the records in question and how easy it is to export the records from this system.

Organisations manage their born digital records in a variety of ways depending on their needs and business contexts. Records may be stored locally on laptops, in shared drives, Electronic Document and Records Management Systems (EDRMS), so called "line of business" systems (such as Case or Contact Management, HR or Finance systems) or even complex eDiscovery systems.

NRS will be preserving the records themselves and not the system they are currently stored in, so some preparation may be required to anticipate this by the depositor, often working in conjunction with their IT department.

Whatever the case, a number of factors in the record transfer process will be determined by the system in use, for example: the structure of files and folders, what metadata is available, document naming conventions and versioning, and restrictions on access to documents. NRS will provide guidance on this throughout the process.

3.2 Redacted Records

There are many reasons that a record may exist in a redacted form in the original record keeping system, but it is very unlikely that any sensitivity will exist indefinitely. **NRS will normally accept redacted records only so long as the original, un-redacted records are also included with the deposit**. This is so NRS can provide access to material once the sensitivity has lapsed.

Both the DRU and your client manager should be made aware of any redacted material ahead of transfer.

3.3 Understanding file formats

A file format is how we describe the way information in a digital file is encoded and made readable by a computer. Generally speaking file formats that are ‘open’ (such as xml or CSV) are more straightforward to preserve than ‘proprietary’ file formats that were created to be used with specific software. It will be helpful for us to know at the outset if your organisation routinely creates files within proprietary systems, or which may otherwise have characteristics which present a challenge to preservation. The DRU will ask what types of records you will be depositing and discuss any potential preservation risks with you.

3.4 File format Restrictions

NRS will endeavour to accept files in the format they were originally created in and used. This is to maintain the archival integrity of the records. **Please do not pre-emptively transform digital records you plan to transfer to us into any other format after their creation and use, without discussing this with us first**. Notwithstanding this there are a small amount of file formats that we do not accept. These are:

- Any compressed files such as zip files. These must be uncompressed ahead of transfer.
- Executable program files.
- Files with any encryption or password protection. This encryption or passwords must be removed ahead of transfer.
- Certain Database files (such as Microsoft Access files).

4 Preparation for transfer

NRS has some requirements that must be met before transfers can be accepted to the digital archive. These preparatory steps are essential to ensure that records can be accessed and understood by future generations.

4.1 Description

It is essential that we receive records with appropriate contextual information that allows us to access and interpret the records. In the short term, this will allow us to confirm what we have received is exactly what we expected to receive from the depositor. In the longer term, this will assist us in cataloguing and help us create a more complete and reliable archival record for future users.

In order that NRS has all required information ahead of transfer, we require you to create a 'manifest' list of what is being transferred – ideally as a CSV or PSV file. This manifest must be submitted ahead of transfer so that metadata can be verified by NRS. We recommend that the program DROID is used to compile this data. DROID has been developed by The National Archives (UK) specifically to obtain essential metadata for digital preservation.

Guidance for using DROID is attached in an appendix to this document, and well as guidance on how to compile this data if the use of DROID is not possible. The DRU are happy to provide further advice and support if required.

The minimum metadata fields we would expect to receive prior to transfer are:

Mandatory Fields	Description	Generated by DROID?
File Pathway	Original file pathway of record	Yes
File Name	Original file name of record	Yes
Size	File or folder size in gigabytes (GB)	Yes
Type	Folder or File	Yes
Date Last Modified	Last modified by user (not system generated).	Yes
Checksum	Computer generated "digital fingerprint". Used to check integrity of files.	Yes
Estimated Date of Creation	Date in format DD/MM/YYYY	No - Manual Input Required
Closure Status	Whether closed or open to access requests upon receipt by NRS. Your Client Manager will provide guidance on this.	No - Manual Input Required
Rights	Copyright and any other intellectual property rights conditions.	No - Manual Input Required
Optional Fields	**Description**	**Generated by DROID?**
Description	Any comments or information from you that provides useful contextual information. E.g. contextual overview of collection, business process	No - Manual Input Required

	description, roles and responsibilities of staff where appropriate etc.	
Identifier	If you have a unique ID reference for the files or folders, this should be included here (this field is mandatory if identifiers exist).	No - Manual Input Required

4.2 Additional Metadata

In addition to the manifest list, we ask that any applicable metadata accompany the transfer of the records. This may include system generated technical metadata files or any indexes or keys that interpret the records. The format of this can vary considerably depending on which system you store your records, so this should be discussed with the DRU ahead of transfer, to ensure we can interpret this information.

5 Test data

In some cases it may be helpful for us to take a sample of your organisation's data, including a sample of the records themselves and any available metadata. This will allow for us to test the records' validity and viability for preservation. This will be particularly important if your organisation is storing records in complex ERDMS or cloud-based systems, where the export of files needs to be verified. If we feel it is necessary to take test data we will discuss this fully with you at the relevant time.

6 Transfer

6.1 Notify NRS

Once your organisation is ready to transfer records to NRS, and you have prepared your data in accordance with section 4 above, please contact your Client Manager at NRS. We will then make arrangements to set a date for transfer. We only accept archival records via encrypted USB drive: we will supply and send you one of these in advance of transfer.

We will ask you to transfer your data on to the drive on a **specified date**, so that we are clear when the transfer happened for administrative purposes.

6.2 Create the transfer package

When transferring your data on to the drive we would ask you to bear in mind the following instructions:

- Put all of your data into **one top level folder** to be named YYYYMMDD[Organisation Name] – for example 20200506ScottishGovernment – please do not use spaces;
- Where possible each transfer should be sent on **one drive** so that transfers are not split into several parts. Contact the DRU on the email address below if this is not possible.

- Do not put any further encryption on the transfer drive.

Any metadata that you have agreed to transfer with the records (see section 4.2) should be included on the drive. If this metadata was not stored alongside the records in the original system, they should also be stored separately on the drive in a folder named YYYYMMDD[Organisation Name]Metadata. A copy of the manifest list described in section 4.1 above should also be included, named YYYYMMDD[Organisation Name]Manifest. Taken as a whole, we refer to the data, metadata and manifest as a 'transfer package'.

If possible, it is advantageous for the transfer to the drive to be conducted using software that verifies the copy using checksums. There are many packages that provide this and your local IT department may be able to recommend one for you. One example of a package that provides this level of verification is Teracopy. Instructions for using Teracopy are listed in Appendix 3 to this document. If using a piece of software such as this is not possible, please transfer using the Windows Explorer 'copy and paste' process, and not 'cut and paste', so you can maintain a copy of the records on your own systems while we process the transfer package, as outlined below. If further assistance is required in the transfer process, please contact the DRU.

6.3 Receipt of the transfer package by NRS

Once we have received the transfer package, it will be quarantined for a period of four weeks to allow us to check it for malware. We will then validate the file formats present and check the data for completeness. Once these processes are complete and the data has been transferred to our digital repository, we will confirm to your organisation that transfer has been successful.

This whole process normally takes 6 weeks from initial receipt. **You must not delete copies of the transferred records until we send this confirmation.** If retaining your data for this period of time is likely to be difficult please let us know in advance so that we can make suitable arrangements.

Should any part of the transfer package fail any of the completeness, quarantine or validation checks described above, processing will be stopped and a request will be made to the depositor for a new transfer package to be sent.

7 After transfer

7.1 Deletion of records

Records should, in most circumstances, be transferred to the archive at a point in their lifecycle when they are no longer required for ongoing business use. It is important the version we store in our digital repository is the final version and there is no risk of the record being brought back into current business use and changed. It is therefore important that your organisation ensure that any copies of transferred records are managed appropriately to prevent this from happening, usually by secure deletion.

7.2 Access to archival records

NRS can provide access to born digital records by either providing copies of files on encrypted drives or by use of the access desk located at West Register House (by prior appointment only). The DRU can conduct a search for records on a user's behalf. The more information we are provided with, the more successful this process will be, so we ask you to maintain a record of what was sent to NRS. Please note that access to records will only be permitted to third parties (other than the depositing organisation) if the records are: classed as open, are not exempt under FOISA, and there are no data protection considerations that prevent access.

8 Further help

If you have any comments or queries on this guidance please contact the DRU at digital_records@nrscotland.gov.uk.

Further information on the management of born digital records can be found on our website at http://www.nrscotland.gov.uk/record-keeping/electronic-records-management.

9 Glossary

Born digital	Records that were created and used in a digital format.
Checksum	Computer generated sequence of letters and numbers that can be used to check data for errors.
Client Manager	Key point of contact from NRS in arranging deposit of records.
CSV	Comma Separated Value – a type of file that can be created easily in excel or other spreadsheet packages.
Digital Preservation	A series of processes and procedures which allow digital files to be kept and made accessible over time.
Encryption	A method of encoding digital files so that they are only readable by those with the relevant authorisation.
Fixity	Checks and evidence that a digital file has not changed over time, for example through unauthorised access or corruption.
Lifecycle	The basic lifecycle of a record, whether digital or paper, is creation, immediate business use, semi-current (mainly reference use in ongoing business), review and disposal. Selection for permanent preservation in an archive is one disposal option.

Metadata	Descriptive, technical and contextual information about a digital file or series of files.

May 2020

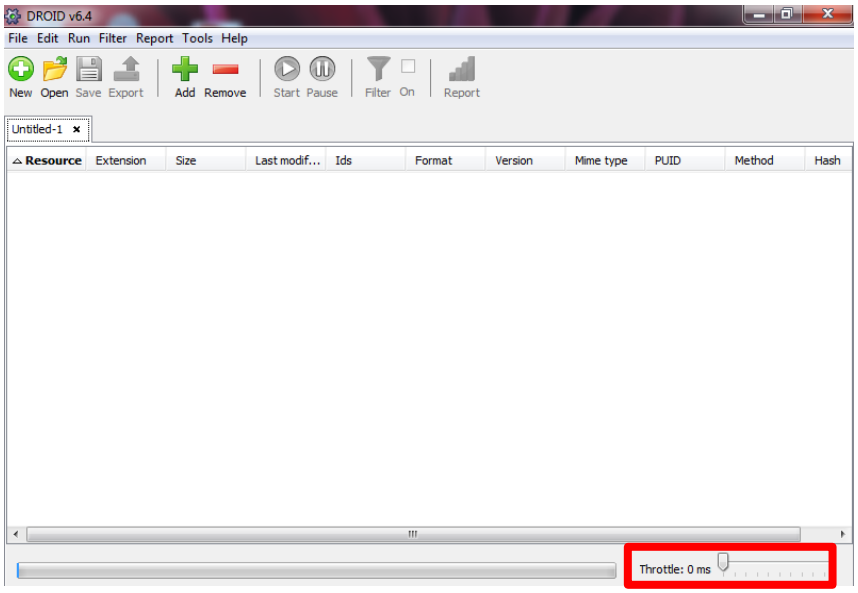
10. Appendix 1: Preparing Manifest Using DROID

DROID can be installed from The National Archives website: https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/

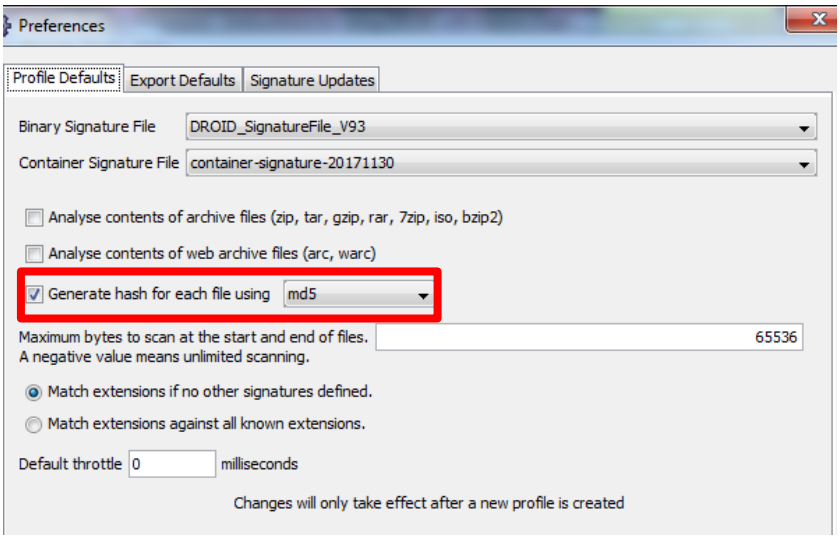
Please refer to this page for installation instructions. In the first instance, always contact your local IT department before you attempt to install DROID. Most corporate IT departments prevent normal users form installing software and you will likely need their assistance. If you are still unable to install DROID, please contact digital_records@nrscotland.gov.uk

Applying correct settings to DROID:

Upon opening DROID, check that the Throttle is set to 0.



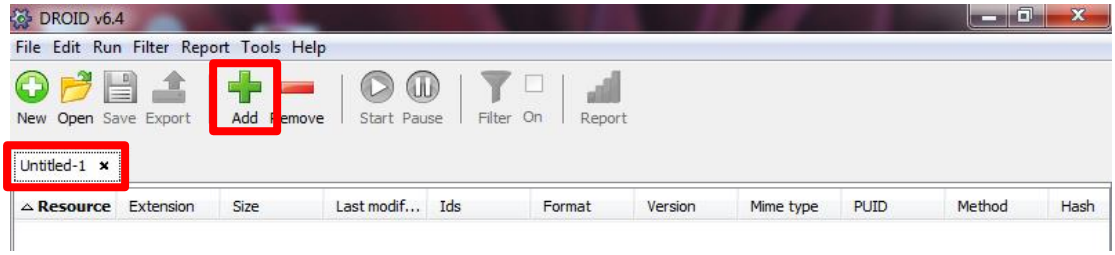
Click on 'Tools', then 'Preferences' and ensure the settings are as below.



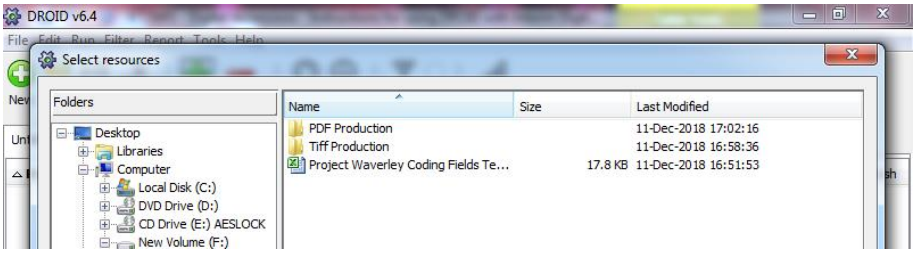
The DROID signature file and container signature file may be different from this screenshot as these are update files for the index DROID uses to identify file formats. The DRU will inform you what file versions these should be.

Creating Profile:

The first time you open DROID, a tab entitled 'Untitled-1' should show on the workspace. If it is not showing click 'New'. To select the folder you wish to profile, click "Add".

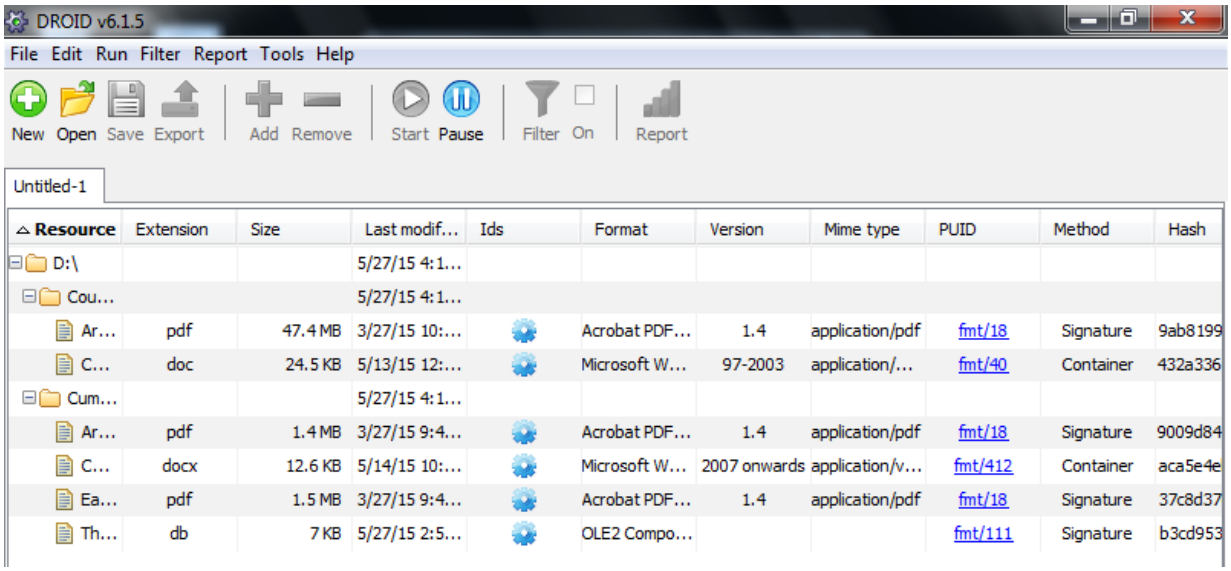


In the left hand pane of the "Select resources" window navigate to where the files you wish to profile are located. Click on the required folder and then click OK.



Select 'Start' from the top bar. A blue bar along the bottom of the screen will show progress. This may take several minutes depending on the size of the folder.

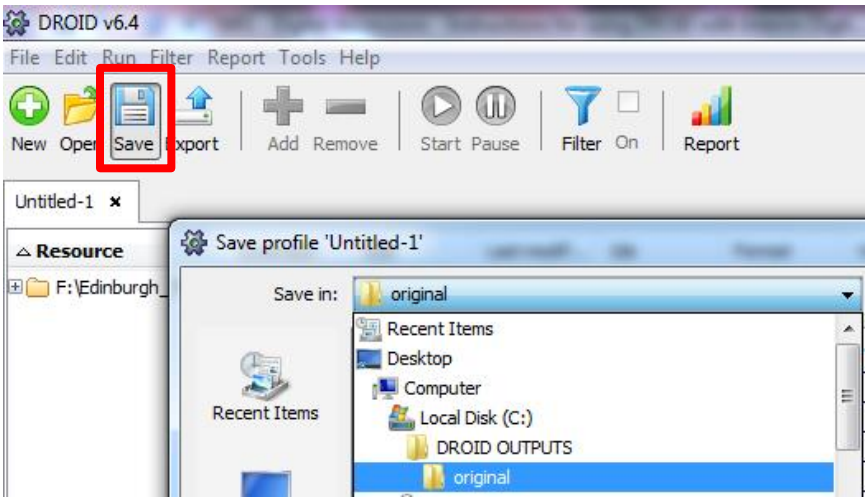
Once complete, if you click on the folder you will be able to see the results.



Resource	Extension	Size	Last modif...	Ids	Format	Version	Mime type	PUID	Method	Hash
D:\			5/27/15 4:1...							
Cou...			5/27/15 4:1...							
Ar...	pdf	47.4 MB	3/27/15 10:...		Acrobat PDF...	1.4	application/pdf	fmt/18	Signature	9ab8199
C...	doc	24.5 KB	5/13/15 12:...		Microsoft W...	97-2003	application/...	fmt/40	Container	432a336
Cum...			5/27/15 4:1...							
Ar...	pdf	1.4 MB	3/27/15 9:4...		Acrobat PDF...	1.4	application/pdf	fmt/18	Signature	9009d84
C...	docx	12.6 KB	5/14/15 10:...		Microsoft W...	2007 onwards	application/v...	fmt/412	Container	aca5e4e
Ea...	pdf	1.5 MB	3/27/15 9:4...		Acrobat PDF...	1.4	application/pdf	fmt/18	Signature	37c8d37
Th...	db	7 KB	5/27/15 2:5...		OLE2 Compo...			fmt/111	Signature	b3cd953

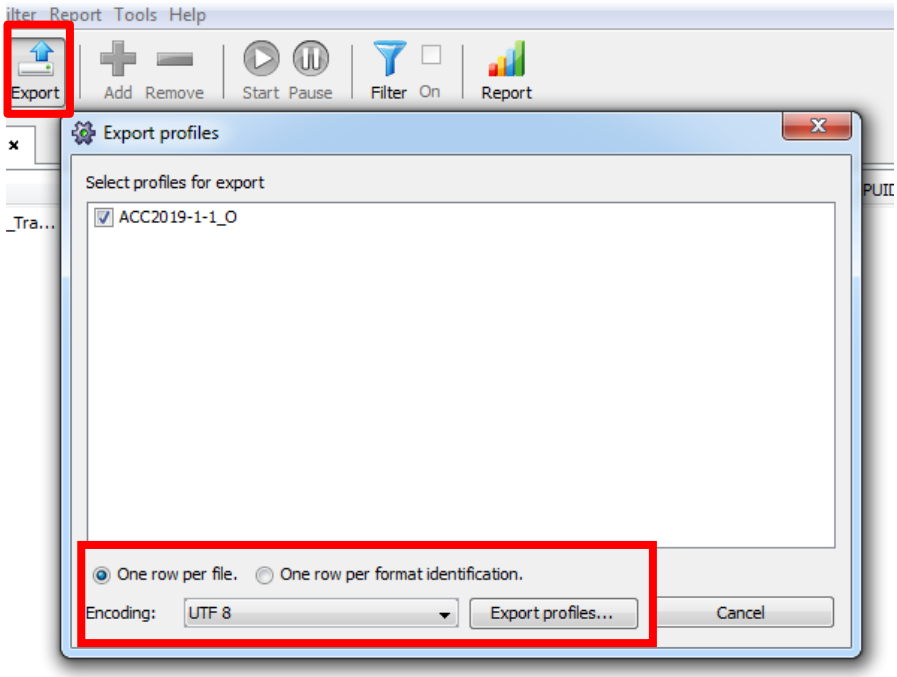
Saving Profile:

To save the profile, select the 'Save' icon and select folder to save the file. Please retain the '.droid' file as you will need to send this file along with the CSV file to NRS.

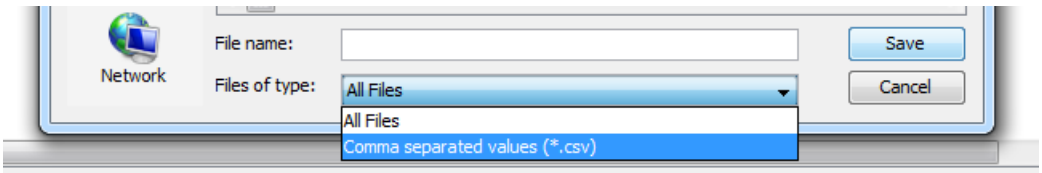


Exporting Profile as CSV:

To export the profile as a CSV file, select the 'Export' icon and ensure the settings are applied as below.



Ensure the '.csv' option is selected from the drop down menu and save the file using the naming convention 'YYYYMMDD[Organisation Name]Manifest'.



Manually Adding Mandatory Metadata Columns to Manifest:

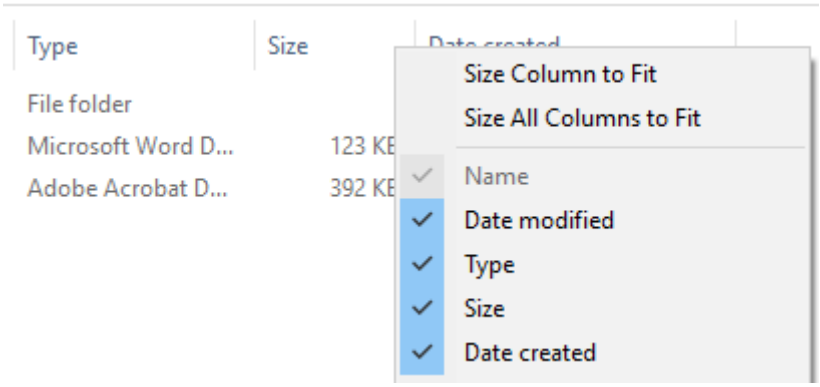
Open the CSV file and add 5 additional columns from column S to accommodate the additional mandatory fields outlined on page 6 of the guidance document. You will not be required to amend any of the fields/columns generated by DROID in columns A-R.

P	Q	R	S	T	U	V	W
MIME_TYPE	FORMAT_NAME	FORMAT_VERSION	Date of Creation	Closure Status	Rights	Description	Identifier
application/msword	Microsoft Word Document	97-2003					
application/pdf	Acrobat PDF 1.3 - Portable Document Format	1.3					
application/rtf, text/rtf	Rich Text Format	1.7					
application/msword	Microsoft Word Document	97-2003					
application/pdf	Acrobat PDF 1.3 - Portable Document Format	1.3					

Estimated Date of Creation:

You can use an excel query to help to populate the 'Estimated Date of Creation' field. Instructions for this is noted in 'Using Excel 'From Folder' Query' below.

You may note that this query only picks up the files themselves and not the folders, so you will need to look up each folder in Windows Explorer to identify the created date. You may need to right click the top bar to ensure the 'Date Created' field is visible.



Remaining Metadata Columns:

The remaining columns should be populated as outlined on page 6 of the depositor guidelines. Your Client Manager can advise on 'Closure Status' and 'Rights' if this is not apparent. 'Description' and 'Identifier' are optional fields but recommended.

11. Appendix 2: Preparing Manifest When Using DROID Is Not Possible

We appreciate there may be some instances where use of DROID is not possible. We recommend following the below steps to put together the manifest manually. As with compiling any data manually, there comes a risk of user error. We, therefore, recommend you compile this data a folder at a time and manually check it is correct as you go along.

The Digital Records Unit will supply you with a csv template with the fields you should populate.

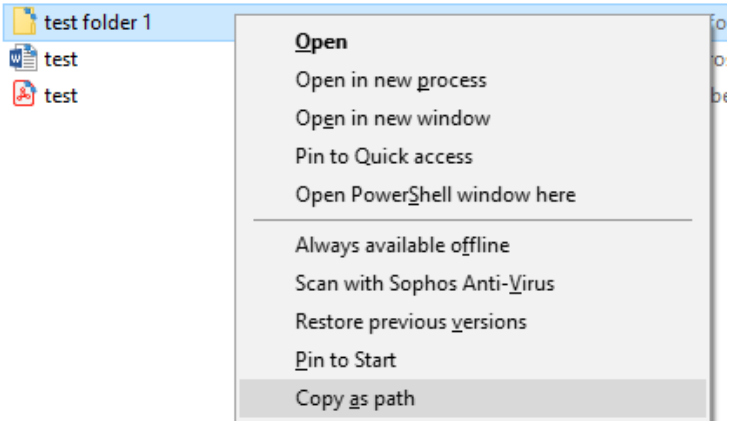
A	B	C	D	E	F	G	H	I	J	K
File Pathway	File Name	Size	Type	Date Last Modified	Checksum	Date of Creation	Closure Status	Rights	Description	Identifier

Using Windows Explorer:

The easiest way to view and analyse the metadata for a file or folder is to view via Windows Explorer.

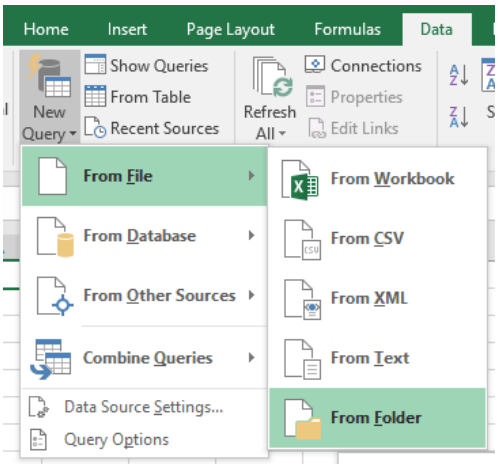
Name	Date modified	Type	Size	Date created
test folder 1	13/05/2020 12:22	File folder		13/05/2020 12:21
test	22/04/2020 09:40	Microsoft Word Document	123 KB	13/05/2020 11:55
test	22/04/2020 09:24	Adobe Acrobat Document	392 KB	13/05/2020 11:55

You can use Windows Explorer to manually check any data you are inputting. To copy the pathway for any file or folder, you would need to press shift and right click, as below. This can then be pasted into your manifest. This can be used to populate data for folders that will be missing when using the excel query.

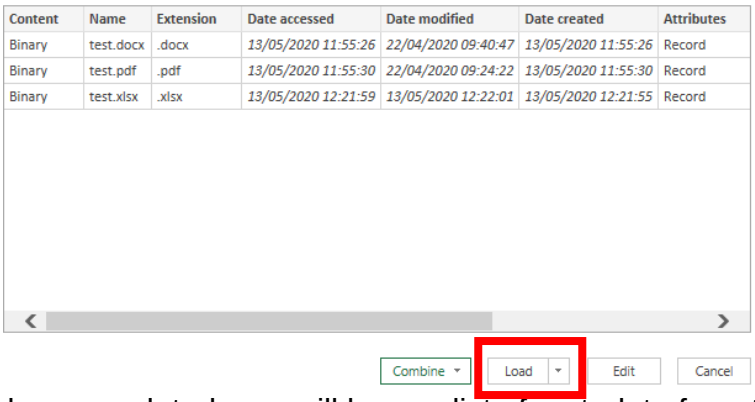


Using Excel 'From Folder' Query

Excel provides a way of capturing metadata from folders, minimising the amount of manual input needed. In the 'Data' tab select 'New Query', 'From File' and then 'From Folder'.



Browse to the folder you wish to profile and select 'Load' as below.



Once the query has completed, you will have a list of metadata from the folder.

Name	Extension	Date accessed	Date modified	Date created
test.docx	.docx	13/05/2020 11:55	22/04/2020 09:40	13/05/2020 11:55
test.pdf	.pdf	13/05/2020 11:55	22/04/2020 09:24	13/05/2020 11:55
test.xlsx	.xlsx	13/05/2020 12:21	13/05/2020 12:22	13/05/2020 12:21

Paste the data from this sheet into the manifest template. You can fill the File Name, Date Last Modified and Estimated Date of Creation field. Anything that is generated by this query will be a file so you can populate the 'Type' column noting this. You will need to manually check for folders using Windows Explorer and populate this data yourself.

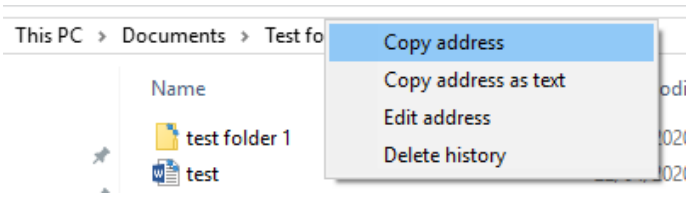
You can also use the 'Folder Pathway' to generate the File Pathway by using the "&" excel function. It's just a case of merging the Folder Pathway with the File Name.

Folder Pathway	File Name	File Pathway
\\TEST\Test folder\	test.docx	=A2&B2
\\TEST\Test folder\	test.pdf	\\TEST\Test folder\test.pdf
\\TEST\Test folder\test folder 1\	test.xlsx	\\TEST\Test folder\test folder 1\test.xlsx

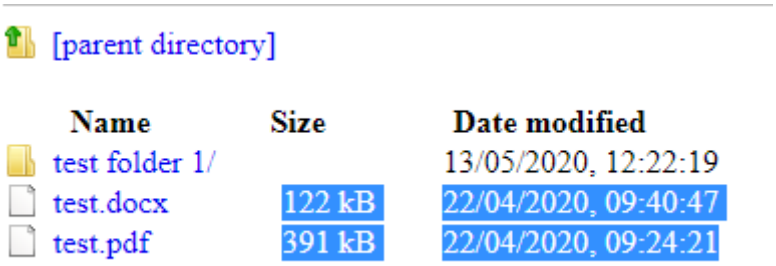
You can then copy the merged cells and paste 'values only' to produce the File Pathway and you can delete the Folder Pathway column.

Using Web Browser to generate size fields:

You can use your web browser to copy and paste the file sizes into the manifest. Look up your folder in Windows Explorer and right click the folder to copy address.



Paste this address into any web browser and this will generate the folder as html.



Simply highlight the Size values and copy and paste this into the Size column on the manifest.

Checksums:

Checksums are slightly more complicated as they generally require a piece of software to generate. It would be worthwhile, in the first instance, to contact your local IT department as they may already have a system in place to generate these. MD5 is our preferred version of checksum so you would need to inform them of this.

If you are running Windows 10, you can use the Powershell to generate the checksums using a simple command. Firstly 'Copy Address' for the folder you wish to generate the checksums using Windows Explorer as described above. Then open Powershell and paste in the below:

Get-FileHash **Paste in Pathway Here*** -Algorithm MD5

The '*' is a wildcard so it will generate a checksum for any files within the folder. You can then copy and paste the checksums from the Powershell into the manifest. Folders do not generate checksums so you can leave these cells blank.

You can also use other tools such as Teracopy to generate checksums. Instructions for this are listed below.

Remaining Metadata Columns:

The remaining columns should be populated as outlined on page 6 of the depositor guidelines. Your Client Manager can advise on 'Closure Status' and 'Rights' if this is not apparent. 'Description' and 'Identifier' are optional fields but recommended.

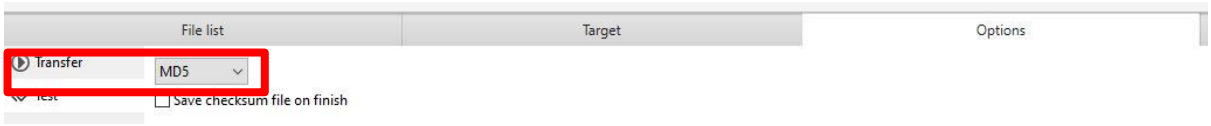
12. Appendix 3: Using Teracopy to Copy Files and Generate Checksums

Teracopy is a piece of software that copies files from location to location and generates checksums to verify the transfer. While there are other utilities that can provide a similar service, such as the inbuilt Robocopy in Windows, Teracopy is one easy to use option. This can be used when transferring the deposit package to the encrypted hard drive and to generate checksums for the manifest, if you are unable to use DROID.

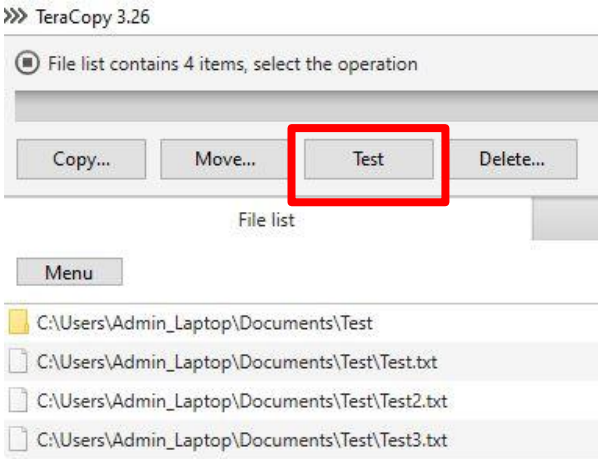
It can be installed for free from: https://www.codesector.com/teracopy

To transfer using Teracopy, right click on folder to be copied and select 'Teracopy'.

Maximise the window and click on the 'Options' tab. Select MD5 from the drop down menu.



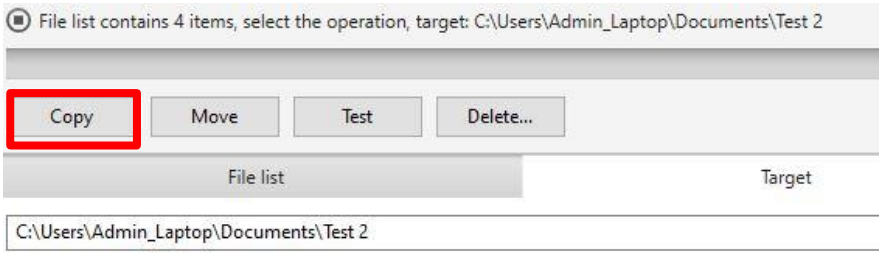
Click on the 'File List' tab. There should be a list of all of the files to be transferred. Click on the 'Test' icon. This will generate checksums for the original folder.



The 'Test' icon will then change to a 'Save Hash' icon. Click on this and this will save the checksums as a '.md5' file. This can be opened in notepad to view the checksums. Save this file as YYYYMMDD[Organisation Name]TeracopyOriginal.

```
1  ; MD5 checksums generated by TeraCopy
2  ; teracopy.com
3
4  47BCE5C74F589F4867DBD57E9CA9F808 *Test\Test.txt
5  08F8E0260C64418510CEFB2B06EEE5CD *Test\Test2.txt
6  D41D8CD98F00B204E9800998ECF8427E *Test\Test3.txt
```

Now click on the ‘Target’ tab and browse to the location you wish to transfer to (the encrypted hard drive). Select the ‘Copy’ icon to transfer.



Once copied, select the ‘Verify’ icon, wait until the checksums are generated and then select ‘Save Hash’. Save the second ‘.md5’ file as YYYYMMDD[Organisation Name]TeracopyTransfer.

If using Teracopy to generate checksums for the manifest, use the ‘Original’ checksums to populate this.

Include both of the ‘.md5’ files on the hard drive when arranging the transfer to NRS. The checksums should be the same between the original and the transfer, as if they are different, the files have been somehow changed during the transfer process. The Digital Records Unit will check these upon receipt and if there are any errors, will return the drive to you so the transfer can be completed again.