

# pdf-title

## Table of contents

<b>1</b>	<b>Tool Demo: DROID</b>	<b>1</b>
1.1	Introduction . . . . .	1
1.2	How does DROID work? . . . . .	1
1.3	DROID and PRONOM . . . . .	2
1.4	Why Use DROID? . . . . .	2
1.5	What information can DROID provide? . . . . .	3
1.6	Downloading, Installing, Opening and Setting-Up DROID . . . . .	3
1.7	Creating a DROID Profile . . . . .	4
1.8	Running DROID and Exporting Data . . . . .	4
1.9	DROID Reports . . . . .	4

## 1 Tool Demo: DROID

### 1.1 Introduction

On this page we will look at the characterization tool DROID, including how to install, open, and use the tool. DROID was developed by the Digital Preservation department of The National Archives (UK) (TNA) and it is free to [download from TNA's website](#) and is supported by a thorough user guide. There is an active community of DROID users and the tool has also been integrated into a number of larger repository systems.

### 1.2 How does DROID work?

DROID uses three different methods for identifying file formats, these are an analysis of the file's:

- Extension
- Signature, and
- Container

The metadata provided by DROID for each file will say which method was used to identify the individual files.

If DROID identifies the file by its extension this means that the format was identified purely on the basis of its file extension. Such an identification may not be reliable, as file extensions can easily be changed, and many file formats and versions of file formats have the same extension.

A signature identification means that a format was identified by finding a specific pattern in the byte sequence, usually in the header of the file. The sequence is unique to a particular file format and version. This method is much more reliable than identification by extension only.

A container identification means that a format was identified by finding embedded files, often with signatures of their own, inside the main file. For example, OpenDocument word processing files are actually zip files containing xml files, images or other resources used in the document. A container identification would identify the main file as an OpenDocument file, not a zip file. This method is very reliable, as not only does the broad type of container have to be identified (for example zip), but the zip file must then be opened, and files inside scanned for further identifications to be made.

### **1.3 DROID and PRONOM**

To allow DROID to make its identifications it needs access to information about file formats and their characteristics to use for comparison. For this, DROID uses PRONOM, a technical registry also developed and maintained by the UK National Archives.

PRONOM is a large database of information on file formats and the software products that support them. A PRONOM record can include information such as the version of the format, what compression and encoding standards are used, if a specification can be accessed and where it can be found, and who owns or manages the specification. This is all information that can be very useful for digital preservation.

When generating metadata about files it has analysed, DROID not only lists the method of identification used, but also includes a unique identifier for the format in PRONOM. This allows a link to be made between the metadata and the corresponding format record in PRONOM.

PRONOM records for commonly used file formats generally include more detail than rare or niche formats. The UK National Archives does, however, welcome contributions from the community to help enrich the data held in PRONOM.

### **1.4 Why Use DROID?**

DROID is a great tool to use when starting out in digital preservation for several reasons:

- First, it is free to download, easy to access, install and set-up, and there is an excellent user guide.

- Second, it has a straightforward, simple user interface. Many other characterization tools do not have a user interface at all, working only from command line instructions.
- It also can identify one of the biggest ranges of file formats amongst similar tools, thanks to the richness of the data in PRONOM.
- Next, we can have confidence in its continued support by The UK National Archives, as they use it in their own digital preservation processes.
- Finally, it can produce, with a high-level of reliability, the basic metadata we need for understanding the digital content we have. It will then allow us to export this metadata in different formats including comma separated value files that can be easily used in Excel or uploaded to a database. It can also produce a variety of summary reports.

## 1.5 What information can DROID provide?

A DROID analysis can produce up to eighteen pieces of metadata for each file. This includes:

- The file name, size, last modified date, and file path: this records what the file is called, how big it is, when it was last edited and where the file is stored on the system
- The identification method used and the status of the DROID analysis: letting us know how DROID identified the file and whether the analysis was successful or if there was an error or problem accessing the file
- The file type, extension, format, version and PRONOM identifier: this covers some key characterization information, so we know what type of file it is and what format (including version) it is. An identifier that links to the relevant format description in PRONOM is also included. Allowing access to more information.
- DROID can also generate checksums according to three different standards to facilitate integrity checking.

## 1.6 Downloading, Installing, Opening and Setting-Up DROID

As mentioned above, DROID is available to [download for free from TNA's website](#). The tool will come packaged in a .zip file. To install DROID, all you need to do is make a new folder to contain the files in the .zip file, and then extract the files to this location. To extract the files, right click on the .zip file and click "Extract All", this will then let you select where you wish to save them.

The following video shows how to open DROID and set some initial preferences:

<https://www.youtube.com/embed/CCOuKB3M3go>

## 1.7 Creating a DROID Profile

When using DROID to analyze files we can save our work in a format called a DROID profile. We create a profile then add one or more folders to it that we would like to analyze. In video below we will walk through creating and saving a profile before adding a folder for analysis.

<https://www.youtube.com/embed/9xT74Tpy7bA>

## 1.8 Running DROID and Exporting Data

Now we have a profile set up and folders added, it is time to run DROID to analyze our files. In the video below we will walk through starting the analysis and exporting the results so we can use the data. It is recommended that you save your exported data as a .csv (Comma Separated Value) file as this can be opened in excel or uploaded into a database for use as metadata. In particular this information can be used in a verifiable file manifest. Also, look out for the “Hash” column which contains checksums.

<https://www.youtube.com/embed/DrBpVoID1nE>

## 1.9 DROID Reports

A really useful piece of functionality offered by DROID is the ability to generate a number of summary reports based on its analysis of files. These reports can provide a comprehensive breakdown of the files analyzed, counts of files organized by a variety of different criteria, as well as reports on any unreadable files and folders. In the video below we'll take a walk through the process of generating and saving a report in DROID.

<https://www.youtube.com/embed/3-97-5eCP7A>