

## **Managing Electronic Records: Checksums**

With more and more government departments choosing to keep electronic records long term (over 20 years) let's take a look at a preservation task that is simple and easy to do (and in most cases, free) to ensure those records remain authentic and reliable over time; checksum generation.

## What is a checksum?

Checksums (also referred to as "hashes" or "fixity") are basically a digital fingerprint for each electronic record. It is a unique string of letters and numbers that let you know a file is authentic and remains unchanged over time. There are multiple types of checksums with the most common being MD5 and SHA-1 but these are the types that you should avoid if possible due to them being "cracked" or vulnerable checksums that can be copied or faked by malicious users. The more secure checksums, in general, are those that start with SHA-2, like SHA-256 for example. There are other types of checksum algorithms available, like the more recently created SHA3 algorithm. Consult with your IT department to determine what is best for you.

### How do I create a checksum?

If you are currently using some kind of records management system, then you may not have to worry about generating and tracking checksums because there is a good chance that it already does. If you are using a network filing system, then you will probably need to generate, track, and validate checksums on your own (you can also always check with your IT department to find out if this is something they are already doing in the background). You can do this rather simply, with the use of a free, open-source, piece of software. There are several different programs available but what you really want is something that will both generate *and* validate the checksums making it easier to check for any integrity issues that may come up along the way. Selecting a program that only generates checksums is fine, but you would have to compare all the checksums yourself which would be tedious, or you could also randomly spot check but there would always be the possibility that you could miss files that have changed and not know it. This is a determination that you will have to make for yourself.

#### Where can I get this software?

Some examples of freely available checksum software are, in no particular order, <u>AVPreserve Fixity</u>, <u>HashMyFiles</u>, and <u>DROID</u> (generates only). There are other options available that you can find with a simple Google search. Ultimately, the choice for what program you choose is up to you and your department and since a lot of the options are free, you can always download and try them out to see what works best for you and your needs. What really matters is that this is a very simple, cost effective, way to ensure that your digital data remains reliable over time.

\\DNP\StateArchivesDCArchive\2003\2003\_094\_Senator\_Scott\_Oelslager'

faaf626f878db40c1fa9c538af3b7cc1eccd7c30a565f3607b57d947a4507ebc f4c6df959ad845a77497da43db04af0df811dda5649fb606328ce70479454e17 7c2875c8195a6e0169f10afb480d1c5e183fe17f0b313bd8bfaea4344b1b1373 841ed62d1307b8c145859bebc1dafcea6ec32c30848260fac8d43108f1e339fe 841ed62d1307b8c145859bebc1dafcea6ec32c30848260fac8d43108f1e339fe 97c0584faba85d3a945ab4b9303f37fbba9515f50521327afb18ebfd9a417d18 55852356d51eec2369c2498274e5576d758d05f3e9245f67c3f1fd875d94fd15 Example of SHA-256 checksums



# Managing Electronic Records: Checksums

## I have the software. Now what do I do?

Typically you'll want to run whatever program you choose the first time just to get a baseline output for your records and then every time you run it after that you will be validating that nothing has changed. How often you validate your files is up to you and what you feel is best for your organization (once a month, bi-annually, etc.) but the sooner you know that something has changed, the easier it is to correct. If, for whatever reason, the checksums *don't* match when you run a validation check, then something has changed at the bit-level of the file(s). If this happens, there is no way to just go into the changed file(s) and correct the error. You would need to restore from backup to a previous version that was unaffected (this is another reason why backing up your data is essential for long-term preservation of electronic records).