

Digital Preservation at the National Library

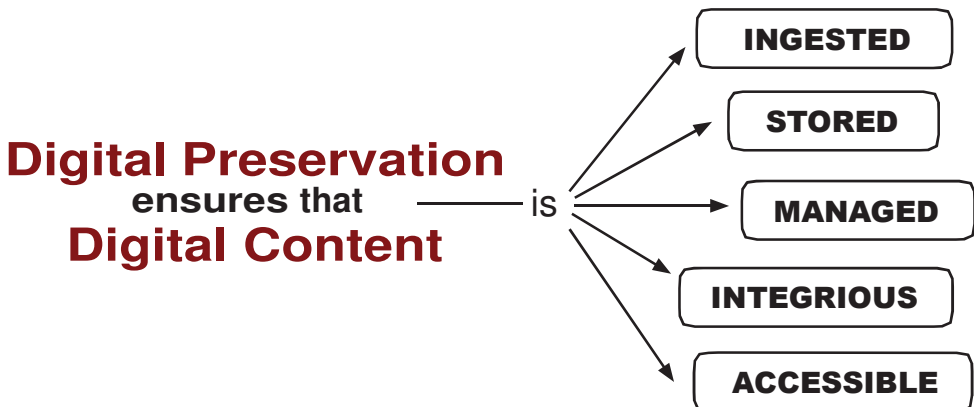
The long-term sustainability of a nation's digital assets is a looming challenge globally.

Obsolete and obscure technologies (hardware, software, file formats etc) threaten our ability to read and understand digital materials into the future.

A famous example is NASA's loss of information about early space exploration because of a lack of managed procedures to ensure the ongoing accessibility and readability of digital records.

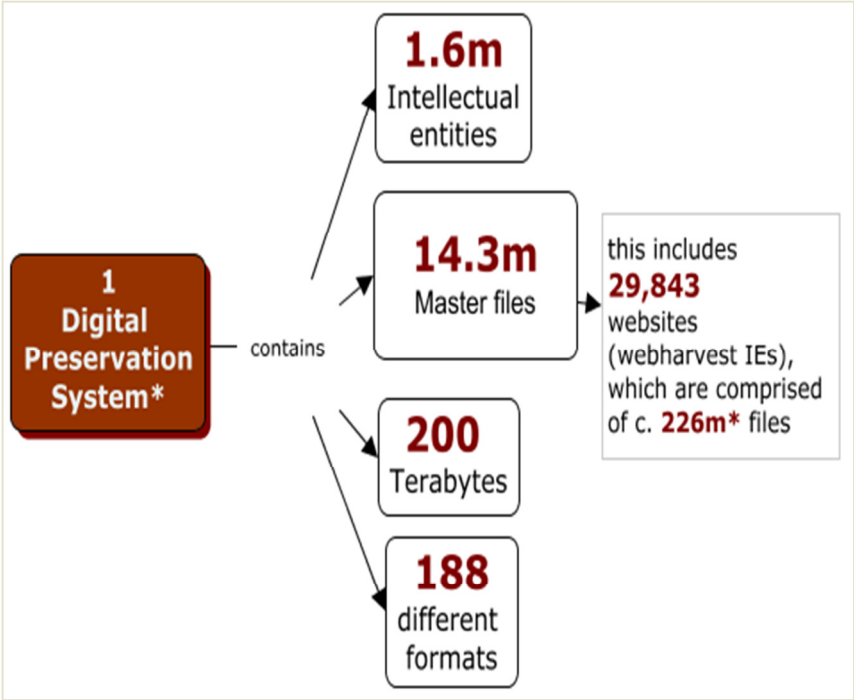
As well as the need for data management the advance in modern technologies also increases the difficulty of maintaining digital materials over time, eg encryption and Digital Rights Management systems.

Our responsibility is to ensure that the digital content in our care is managed and preserved in a way that reflects its status as an asset and taonga of New Zealand.



Our collection

The following diagram shows what our digital preservation system contains (as of 31 May 2017).

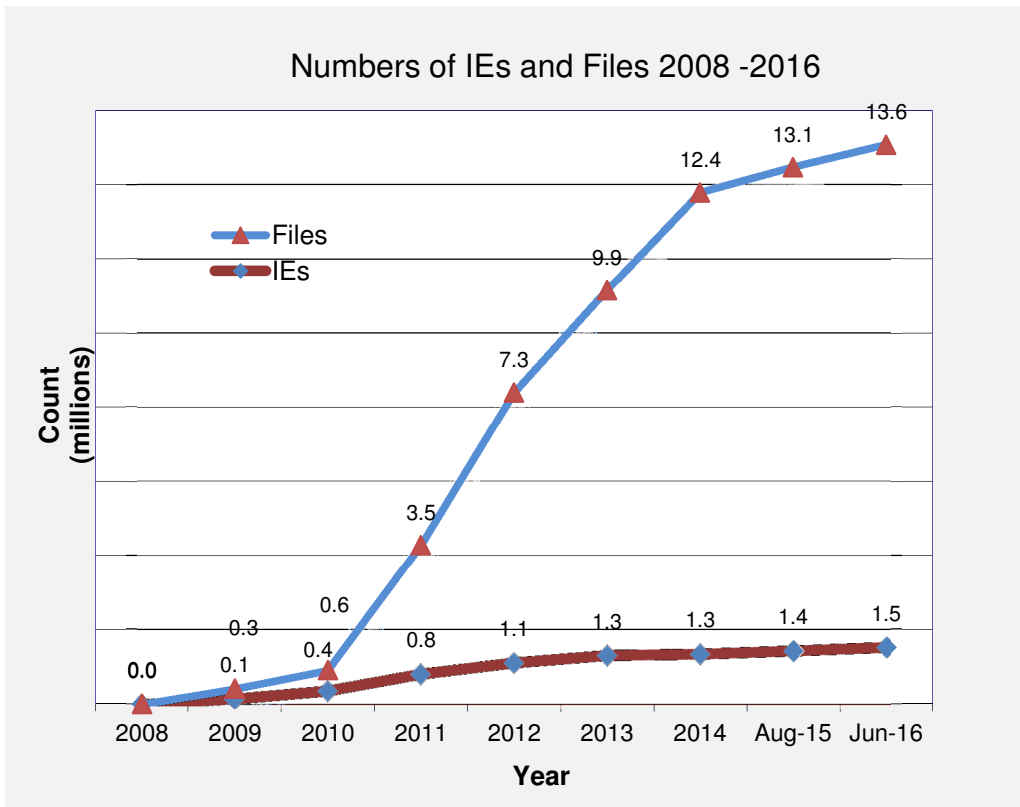


Collection growth

The following graph shows the growth in the number of intellectual entities and files contained in the repository.

Intellectual entities are the objects we preserve (books, albums, photographs, websites), and files are the building blocks of that content.

Since 2010, we have seen a marked increase in the number of files used to make up an object; reflecting an increase in complexity.

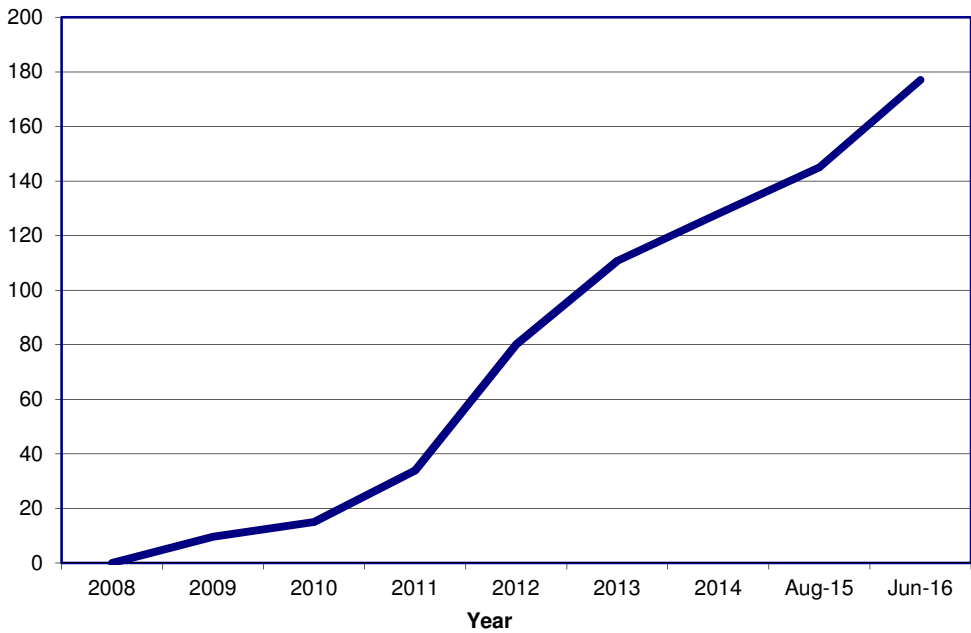


Growing repository

The size of the digital preservation repository has unsurprisingly been growing since inception. From the point of moving into a production environment, the growth has increased at a greater rate almost every year.

Repository growth 2008-2016

Terabytes

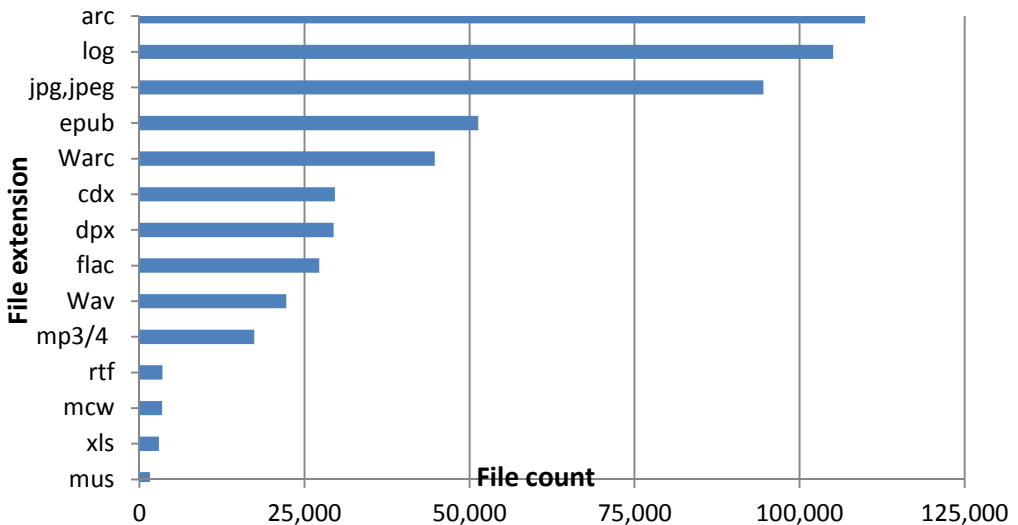
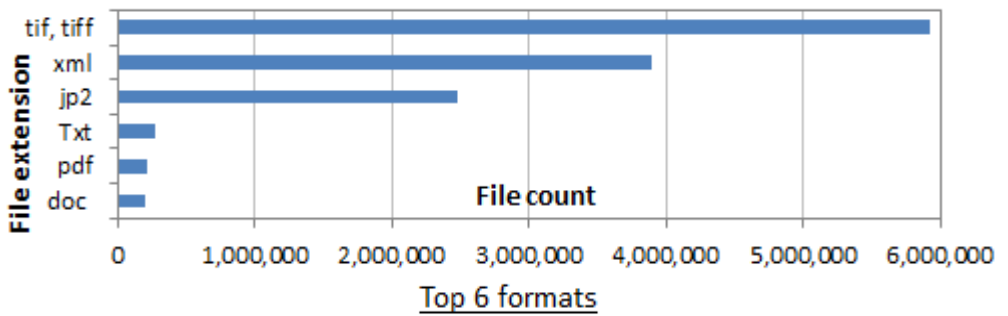


Content format: A key challenge

The format that content is created in provides one of the key challenges for digital preservation: it is this format that can impede access to the content.

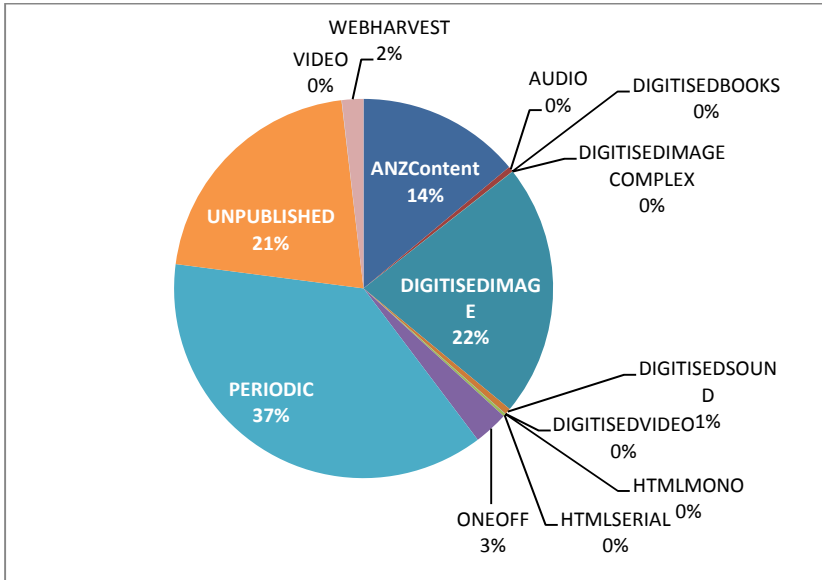
The content in the repository is conveyed in over 188 different formats. The graphs below show the most common formats we currently have.

Top 20 Formats in our collection:

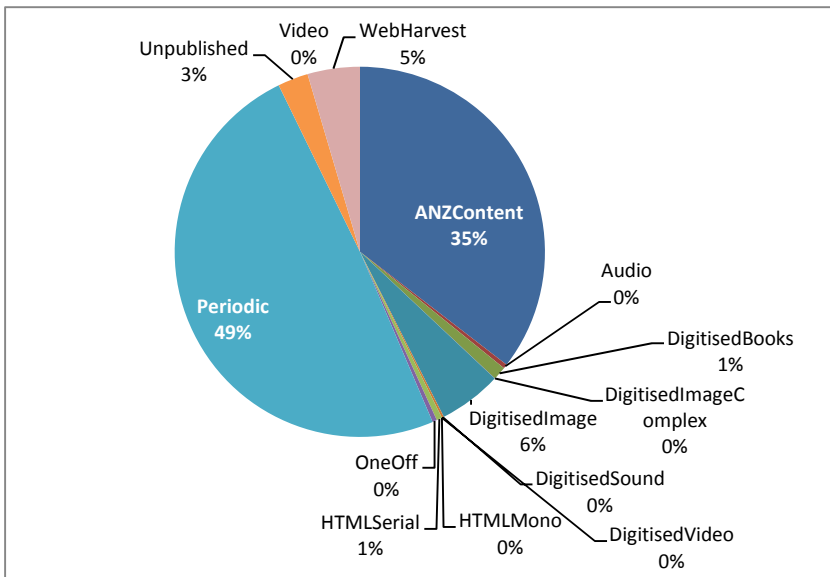


Content type

a) by intellectual entities



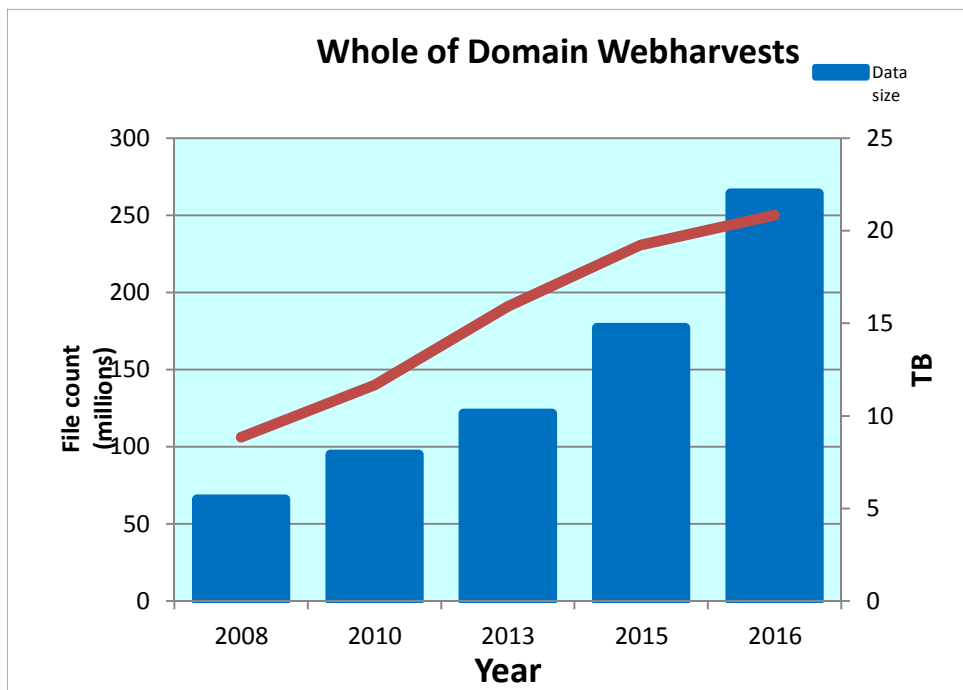
b) by material type (files)



Whole of domain web harvests

The figures contained in the other graphs do not contain the whole of domain web harvests that the Library has carried out.

These harvests collect all **.nz** websites that are publicly accessible. Together, these domain harvests total over 900 million files and 60 terabytes.



For more information visit: <http://digitalpreservation.natlib.govt.nz/>

Contact email: ndha.admin@dia.govt.nz