

Locating Inactive Records on Large Network Drives: Testing Archives Finder

ANTHONY COCCIOLO

Abstract: Recent initiatives in accessioning born-digital archives have focused on removable media, such as using forensic tools to image media. However, there has been little discussion of the born-digital archiving needs of institutional archives. In institutional settings, terabytes of records with permanent value often reside on large, unstructured network drives, often alongside active records. Research at the National Archives of the UK found that up to two-thirds of government information is held on unstructured shared drives with some departments holding up to 190 terabytes of information.

Tools to identify batches of inactive records, such as the records of departed staff members or initiatives that have long ended, are often lacking and are designed more for IT departments to manage disk space. This study will explore some tools for locating batches of inactive records that can then be subjected to appraisal. In particular, it will focus on a script developed by the researcher called *Archives Finder* that aims to address some of the issues with existing tools for locating batches of inactive records. *Archives Finder* searches across large, unstructured network drives for the largest possible grouping of records that are a given number of years old defined by the user. It also includes “fuzzy math” feature that allows the user to specify that only a certain threshold of files need to be X years old. This tool, combined with other tools, will be tested against a large, unstructured network share for its ability to locate inactive records efficiently and accurately.

About the Author:

Anthony Cocciolo is an Associate Professor and the MSLIS Program Coordinator at Pratt Institute School of Information, where his research and teaching are in the archives area. He completed his doctorate from the Communication, Media and Learning Technologies Design program at Teachers College Columbia University, and BS in Computer Science from the University of California, Riverside. You can find out more about him at his website: <http://www.thinkingprojects.org>.