Digital Preservation Archive School

Aims of the course
To provide an overview of the core principles of digital preservation, give attendees the experience of hands on usage of some digital preservation tools and enable them to be able to apply the knowledge gained to their own place of work.

Commitment: 5 days onsite learning at The National Archives over the course of 6 months. Each session will begin at 10am and finish by 5pm.
Homework carried out between sessions in your place of work – it will not be assessed but outcomes will be discussed with the group in following sessions.

Class size: Between 10 – 15 attendees

Class location: The National Archives, Kew, Richmond, Surrey, TW9 4DU

Travel & one night accommodation paid for (if needed)
Lunch provided (please inform us of any dietary requirements)

Course modules
Session 1 – Knowing what you have. Part 1 – 1 November 2019
An overview of digital preservation and practical use of a file format identification tool to run over your digital files and provide information on your digital collections.

Session 2 – Knowing what you have. Part 2 - 29 November 2019
In a continuation of knowing what you have, looking at file format research and how you can get your unidentified formats added to PRONOM and DROID.

Session 3 – Describing what you have – 31 January 2020
How to use tools to extract metadata from your digital files and an overview of the types of metadata validation carried out at The National Archives. A practical session on the use of a metadata extraction tool.

Session 4 – Keeping what you have safe – 28 February 2020
Assessing risk and the different storage options available for digital collections.

Session 5 – Providing Access - 27 March 2020
TBD

Prerequisites: An eagerness to learn the basics of digital preservation and a willingness to work in a supportive and fun group. **No previous digital preservation experience is required.**

Laptops will be provided for the onsite training sessions but to do the homework and allow you to use the skills you've learned after the course has ended **you will need to have installed the following FREE tools on a computer/laptop at work**:

DROID: http://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/ - DROID is a file format identification tool. DROID runs over a folder structure in order to provide file information such as file type, checksum filepath, extension etc.

Java version 1.8 (also called version 8): https://www.java.com/en/download/ - this is required for DROID and CSV Validator to run

Python 3.7.5 https://www.python.org/downloads/ - required to run Python scripts we'll use to demonstrate how to manipulate metadata and utilise some aspects of Apache Tika.

CSV Validator: https://search.maven.org/search?q=g:uk.gov.nationalarchives%20AND%20v:1.2-RC2
Click the download arrow alongside the csv-validator-ui option and select allpication.zip from the download options.
CSV Validator is used by The National Archives to check the integrity of metadata csv files.

Apache Tika: https://tika.apache.org/download.html - Apache Tika can be run over files to extract administrative metadata that can then be preserved with the files.

HexEditor: https://mh-nexus.de/en/downloads.php?product=HxD20 – HexEditor is used to view the byte code of digital files. HexEditor allows you to drag and drop files in order to interrogate their byte sequences for file format research but also to interrogate issues you may be having with a file - it provides a compare function that allows you to find disparities between files that should be identical but are showing different checksums. It allows for a variety of functions but The National Archives use it for file format research and trouble shooting.

TeraCopy: http://www.codesector.com/downloads - this allows you to copy files from one location to another and carry out auto verification checks to ensure the copy was successful. Its focus is data integrity, file transfer reliability and the ability to pause or resume file transfers.

Notepad ++: https://notepad-plus-plus.org/

Gizmo: http://arainia.com/software/gizmo/