Notes for Archiving

Background:

At the end of each academic year, a static copy of the completed courses is required, which is then transferred to the archive server (archive.st-andrews.ac.uk). Previously, HTTrack was used for this purpose. Before transferring to the archive server, various updates were applied to the copied course pages, such as adding a green banner and modifying links and titles to indicate they are archived. These edits were made using the find and replace function in the code editor.

Issue:

The HTTrack tool can no longer be used due to system changes that block its use. This solution utilizes PHP to replicate the HTTrack tool's functionality by:

1. Obtaining a list of course URLs for UG and PG:

a. Note: Run the script separately for each UG & PG.

b. Initially use a test sample of 2-3 files.

2. Looping over these files.

3. Creating a local folder structure matching each URL from the list.

4. Adding an index.html file to each folder.

5. Retrieving the content from each URL.

6. Updating the contents, including adding a green banner and breadcrumb links.

7. Writing the updated files to the index.html files within the new folder structure.

8. Addressing the issue with footer removal after completing the above steps using Find and Replace in the code editor.

9. Checking the files locally to ensure they look fine.

10. Moving these files to the archive server under the "courses" directory and adding a folder for the appropriate year.

a. Note: The URL structure does not match the structure of the current archived content, so manual creation of the folder structure on the server is necessary. Subsequently, the appropriate files should be copied to each of these folders.