HMD Ingest workflow tool

This tool, written in C#, provides functionality for reading values from the HMD sharepoint site and automating the ingest workflow process by performing various operations on the values returned from sharepoint. The tool provides outputs in various formats, in particular by writing to text and csv files but also by writing directly to the sharepoint site.

Users

The tool is supplied as a Windows executable file which can be run with no prior installation.

To run, double-click the executable. You will be presented with a series of yes/no options.

Environment

Two environments exist within the software tool - `test` and `prod`. Using `test` sets up a test-based environment, only looks at the test sharepoint site and writes output files to your desktop. Using `prod` sets up the live production environment and uses the live HMD sharepoint.

WARNING - UNTESTED FEATURE IN `prod` environment
Image order csv generation in `prod` is untested:

- Currently I haven't been able to get the ImageOrder.csv files to write directly to the network location due to insufficient permissions. You may have sufficient permissions and be able to write to this location, however!**
- Not recommended to test this yet, but if you feel like it then testing on single shelfmarks is the most sensible way forward here.
  The above does not apply to the `test` environment.

You will be asked if you would like to use an ingest spreadsheet as an input. Select yes/no, and if you would like to use one then make sure you have this in the same directory as the .exe file and that it is called IngestSpreadsheet.xlsx containing a sheet called "PSIP Generator Fields"

The tool then asks if you would like to search sharepoint using a shelfmark. You can use a list of shelfmarks, in a .txt file called 'shelfmarks.txt', or you can use individual shelfmarks. If you answer 'yes' here, followed by your shelfmark of choice (matching Sharepoint exactly!) to use this feature.

You can also select a `project` to run over. The default project is environment-dependent, but the program will tell you what it is currently set to. To change it, type `yes` (then `Enter`) and type the name of the project to analyse. At the moment this is required to match the value in SharePoint exactly.

Checks

As far as possible, individual checks are separated and the user is free to choose which checks to run over.

Four broad checks are available:

- Shelfmark source folder checks This checks that the value of `Source Folder` given for an item in SharePoint is a valid path on the BL network and writes the outcome of this check to Sharepoint.

- Run shelfmark protected character check Checks a shelfmark as entered in SharePoint for protected characters that will cause errors further down the workflow. Writes outcome of this check to SharePoint.

- Generates image order CSV and perform ALTO XML checks Writes to SharePoint if errors are found. Image order csvs are written to a folder on the desktop currently. Queries Iams if relevant to find foliated child shelfmarks and adds these into the image order csv.

- Query library metadata APIs Looks the item up on Aleph or Iams, depending on the metadata source field entry in sharepoint. Depending on whether the item comes under the remit of Aleph or Iams, writes a CSV to a folder in the the user's desktop. The folder name uses the item shelfmark within the name.

For Developers

Structure of the tool

The tool is controlled with the Program.cs file, with different functionality split across the other `.cs` files in the repository:

- SharepointTools.cs
  - Functions to interact with the HMD sharepoint site
  - Verifies site exists, loads from the digitisation workflow sharepoint list and prints specified fields for each item (shelfmark)
  - Creates HMDObject items for each object returned from the sharepoint query. These have member functions of shelfmark, ID, system number and so on.
  - HMDObjects are used as an input for the other functions defined in the `.cs` files listed below
- InputOrderSpreadsheetTools.cs
  - Contains functions to generate input order CSV file for each shelfmark provided and perform XML checks.
- LibraryAPIs.cs
  - Contains functions to query library metadata APIs, return CSV files and write to sharepoint.
- TextFileOutput.cs
  - Functions to output text files after running some analysis on items in sharepoint, used for development purposes.
- HMDSharepointTests
  - Various test functions for the software tool, mainly concerned with testing access to the sharepoint site.