Converting InDesign files to HTML
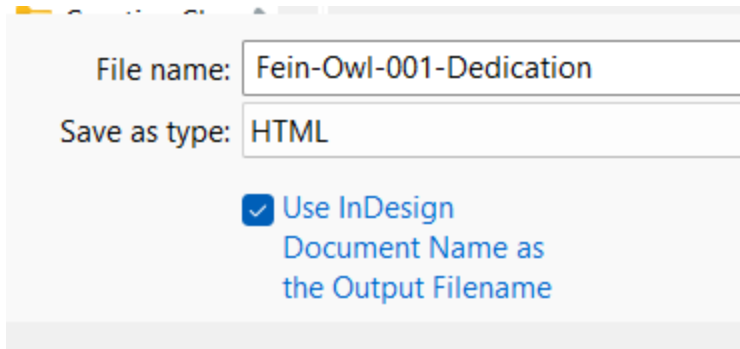
First, we need to convert our InDesign files into HTML. The resulting HTML files will be run through our transformation script to produce our (mostly final) TEI-XML files.

- ❖ Open each InDesign file in the book. Make sure you are working from the Export-XML Book.
- ❖ File > Export
    - ➢ The shortcut keys are Control+E
- ❖ Check that the file type is marked HTML.
    - ➢ Sometimes the window will default to exporting to Adobe PDF Print. If it does, you'll need to click the dropdown menu and select HTML.
- ❖ Select the correct folder to save these files in.
    - ➢ In the Editor's folder, there will be a folder labeled HTML Files. Select that folder.
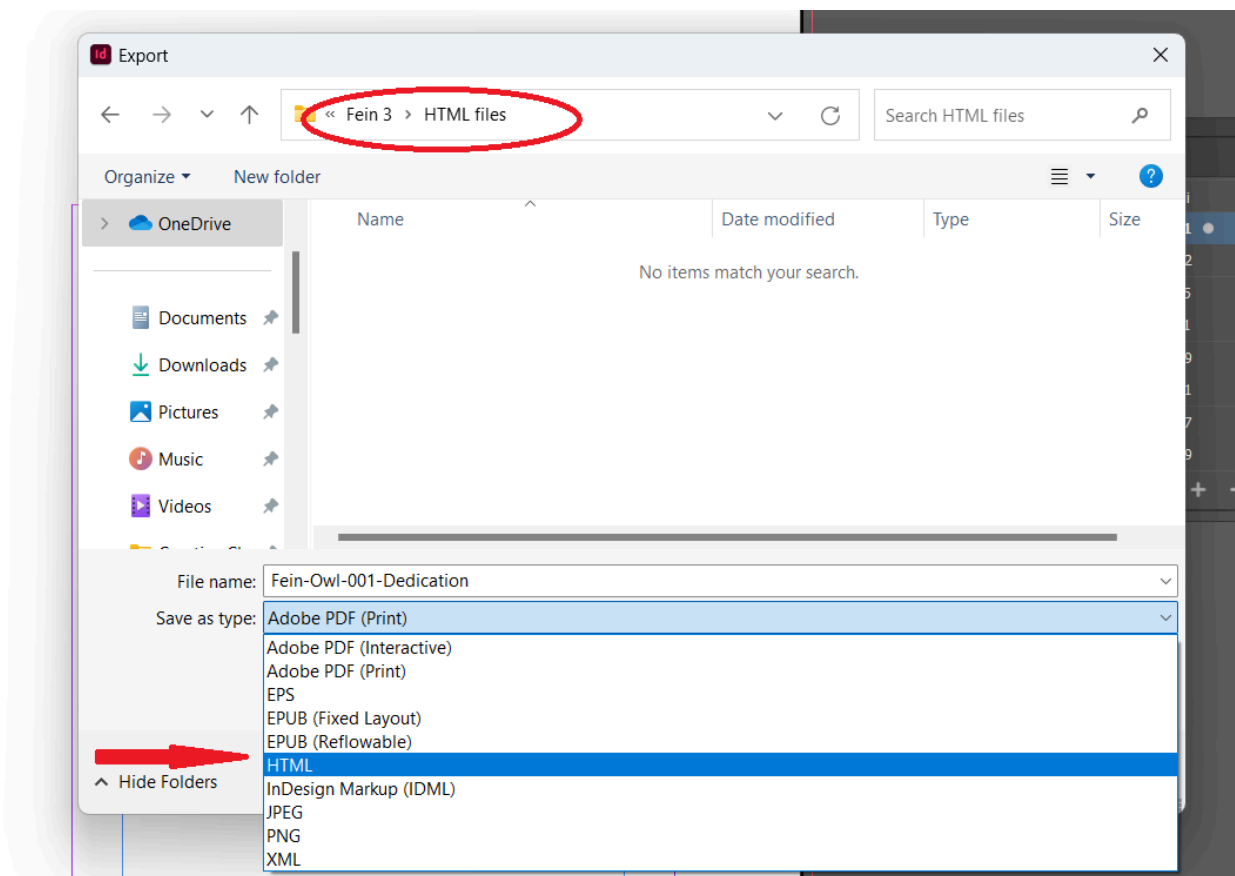
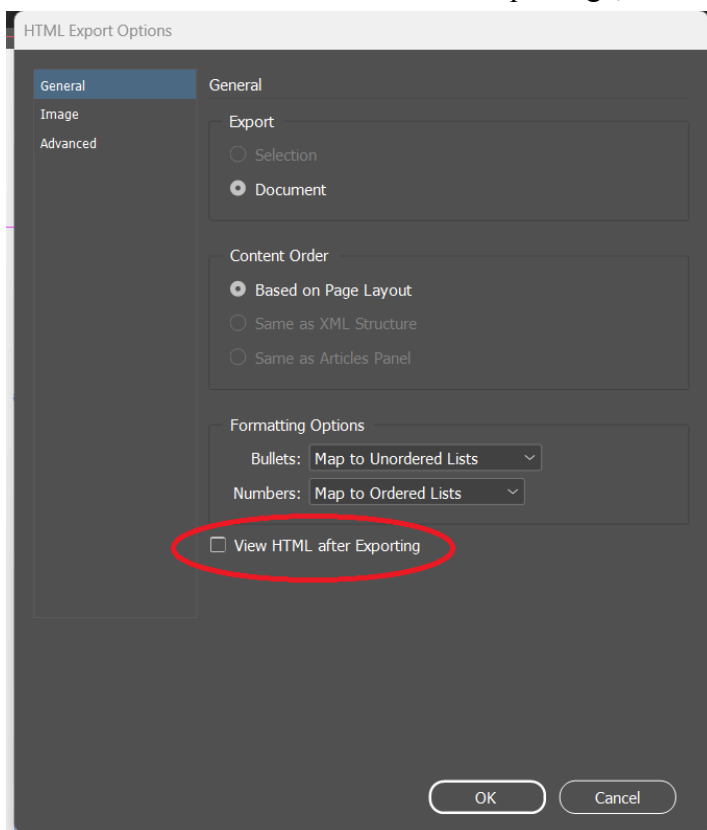- ❖ Make sure the "Use InDesign Document Name" box is checked.

- ❖ Then click save. This will bring up another Options window.
- ❖ Uncheck "View HTML after Exporting", then click Export.

It is important to note that, if you have unchecked the "View HTML after Exporting" option, nothing will open once you have clicked the button that says Export. If you want to make sure that the files have exported properly, you will need to go to HTML folder you saved the files in.
- ❖ If everything worked properly, you will find an html file along with a folder labeled web-resources. This is normal. We don't use the web-resources folder for our export process, so you can delete those.
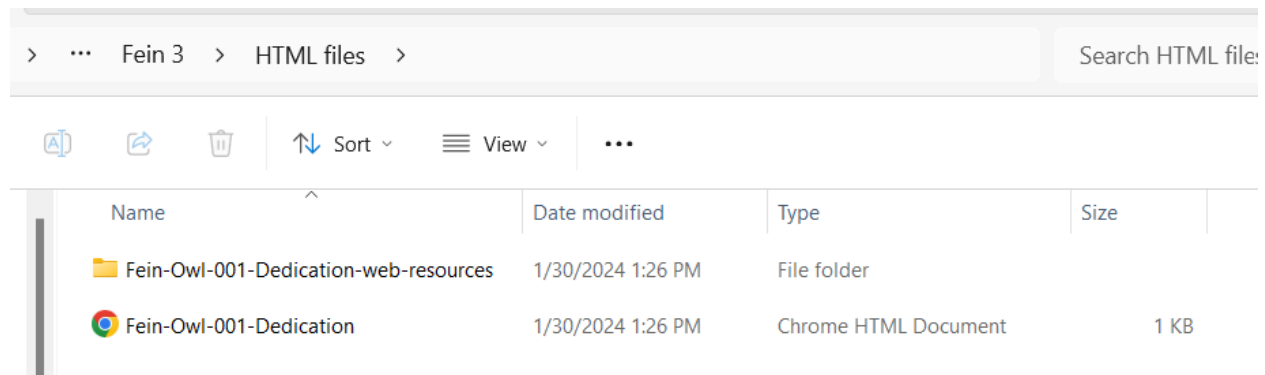
❖ When you have finished creating your section of HTML documents, please upload the HTML documents into the HTML folder for the pertinent edition in Box. Then, upload the corresponding InDesign documents into the InDesign folder the pertinent edition in Box.

Transforming HTML to TEI-XML: The Webform

Once all the HTML files have been transformed, we need to run them through our transformation script to create workable TEI-XML files.

Our script will be posted in our Github. When we are using the script, we will access it through a webform that has been created for us.

The webform has been streamlined as much as possible to prevent duplicated data entry if we have to rerun files multiple times. (This is the reason we have so many spreadsheets at the moment.)

The webform looks like this:

To use the webform, you will need:
1. The HTML file that was exported from InDesign
2. The file's complete title as we want it to display in the TEI Header

Adding titles
You'll need to manually type the title into the form every time you run a file through the webform.
- ❖ If you are working with the Introduction to an edition, the title should be
  - ❖ Introduction
  - ❖ General Introduction
    - ➢ If there are individual introductions or headnotes associated with texts
- ❖ If you are working with an apparatus file specifically linked to a text, you should include one of the following phrases before the title of the text:
  - ❖ Introduction to
  - ❖ Explanatory Notes to
  - ❖ Textual Notes to
- ❖ Any other files' titles should exactly match the title listed on the Table of Contents

Adding HTML files
Either drag and drop a file into the field, or click "Choose File" and select the correct file through the File Explorer.

As of October 2024:
When you click Submit, the webform page will refresh. First on the page will be debugging information, and below should be a link to download the XML file:

================BEGIN DEBUGGING INFO==================

[debugging information]

================END DEBUGGING INFO==================

Download XML File

Download the file and save it in the XML Files folder for the edition you are working on.

Sometimes the transformation script encounters code in the HTML that it is not expecting. If that's the case, then the webform will give one of several responses.
- The debugging information will indicate that one of its processes has failed
- The final "END DEBUGGING INFO" line and the download link will not be on the page.

If this happens:
- First, check the HTML file you loaded in. HTML files that have been formatted and indented don't work well with the transformation script. If it has been formatted and indented, then you may need to re-export the HTML from the InDesign file.
- If this fix doesn't work, then let the Managing Editor know. You'll need to contact our programmer to have him take a look at the InDesign file, the HTML file, and his script.
  - Either he will need to adjust his script, or we will need to adjust our policies in InDesign to fit with the script's existing parameters.