Formatting for files:

Remove Gutenberg text from beginning AND end

Remove all TOC/index/title page info

Remove introductions

Keep chapter headings

melville\_moby\_dick\_pg1234.txt

Keep footnotes unless it’s easy to remove in one go.

Formatting for metadata:

Filename: complete filename as it appears in Box Folder

Title: just write title, no moving “the”

Author: Last Name, First Name; expand if known

Year of Publication: translation year

Genre: eg “Horror”, “Historical”, “Teen wizard”; semicolon-delimited

Form: “novel”, “play”, “collection”

Keywords: semicolon-delimited

Country of origin:

Length: word count? Chapters?

Original language: Specify English if not translated