kwalk
a simple program to crosswalk metadata

Kirsten Vallee

December 1, 2023
You need to upload 1,000 items to your institutional repository from a source like Lens.org or PLOS journals.

You obtain informal metadata for the items on your own or from another person either from creating the spreadsheet from scratch, exporting the data, or web scraping each individual record.
You might need to

**rename**
- Rename all the fields from the invented field names to TIND/your repository platform's field names

**add**
- Add some fields that are missing

**exclude**
- Exclude some fields you don’t want

**combine**
- Combine several fields into one field

**modify**
- Modify the values of date formats or author names in a programmatic way

**generate**
- Generate syntactically correct upload URLs from a simple filename field
What can you do with kwalk?

- Write a simple crosswalk that you can apply to each batch of metadata as you receive it.
- Have multiple crosswalks for multiple projects as you work on them in an intermixed fashion.
- Apply special functions to modify date formats, combine literal and field name text, generate uniform upload URLs, and more.
What do you need to kwalk?

1. A crosswalk in .txt format (i.e. lit-kwalk.txt)

2. A database in .csv format (i.e. lit-data.csv)

*This example is a simple one. There are special functions in the program that can modify date formats, combine literal and field name text, generate uniform upload URLs, and more.*
How to kwalk

01
Open a program like Windows PowerShell

02
Navigate to the directory where your .txt and .csv files are located

03
kwalk-tind -c crosswalk litkwalk.txt litdata.csv > litupload.csv
<table>
<thead>
<tr>
<th>245__a</th>
<th>269__a</th>
<th>7001__a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pride and Prejudice</td>
<td>1813</td>
<td>Austen, Jane</td>
</tr>
<tr>
<td>Brave New World</td>
<td>1932</td>
<td>Huxley, Aldous</td>
</tr>
<tr>
<td>Moby-Dick</td>
<td>1851</td>
<td>Melville, Herman</td>
</tr>
</tbody>
</table>
Thank you

Kirsten Vallee
University of Chicago
vallee@uchicago.edu