OpenRefine

A free, open source, powerful tool for working with messy data

https://openrefine.org/

OpenRefine 3.6.2

Towards a toolbox for the Digital Humanities

https://schoeneh.eu/histotools

I'm no expert in these tools, however:

We have nothing to loose but our inexperience.
We have a whole world to win.

Therefore: Let's have a go!

OpenRefine:

- **Transform**
    - parsing and converting
- **Edit**
    - cleaning, editing, linking and assembling
- **Explore**
    - getting a general idea/overview; finding patterns and systematic errors

data sources

- local file
- URL
- clipboard
- database (SQL)
- Google Drive (sheets)
- [Wikidata/Wikibase]

https://docs.openrefine.org/manual/starting

file formats: input

- TXT
- CSV/TSV
- JSON, XML
- ODS - OpenDocument Spreadsheet
- XLS/XLSX - Excel Spreadsheet
- PX, MARC
- JSON-LD, N3, N-Triple, Turtle, RDF/XML - RDF data
- Wikitext

https://docs.openrefine.org/manual/starting

file formats: output

- CSV/TSV
- JSON
- ODS - OpenDocument Spreadsheet
- XLS/XLSX - Excel Spreadsheet
- Google sheet
- database (SQL)

https://docs.openrefine.org/manual/exporting

use cases (examples):

1. **Text to table/spreadsheet**

2. **geoJSON to table/spreadsheet**

3. **Exploring, cleaning and editing of data**

Example datasets:
https://github.com/Schoeneh/sharing-is-caring/tree/main/Workshops/OpenRefine