From a messy spreadsheet to Wikidata

Start mass data uploads with OpenRefine

Alicia Fagerving
alicia.fagerving@wikimedia.se
Why am I here?
What can I share?
What would I like to learn?
Agenda of the workshop

- Our tools: OpenRefine and QuickStatements
- A quick tour of transformations
- We get some real life data and process it
- What is reconciliation
- Uploading your changes
Why we use OpenRefine

- Quickly get an overview of a dataset, regardless of format
- Easy to estimate how well the topic is covered on Wikidata
- Open source tool with many users – we don’t have to maintain our own tools
OpenRefine

- Open source desktop program
- Runs on a server on your computer and is accessed via a web browser
- Explore datasets in different formats as a table
- Clean and transform data
- Reconcile and match data (identify Wikidata item corresponding to specific entry)
- Upload data to Wikidata

openrefine.org
github.com/OpenRefine/OpenRefine
QuickStatements

tools.wmflabs.org/quickstatements

- Edit multiple Wikidata items
- Create new items or modify existing ones
- Add claims, labels, descriptions, qualifiers and references
- Run in background – you can close your browser window
Uploading your changes to Wikidata

OpenRefine uploader

If same property and value but missing reference, adds reference to existing statement (reference merging).

If same property and target value but different qualifier, creates new statement.
Uploading your changes to Wikidata

OpenRefine uploader

If same property and value but missing reference, adds reference to existing statement (reference merging).

If same property and target value but different qualifier, creates new statement.
Uploading your changes to Wikidata

QuickStatements
Also merges references.

If same property and target value, adds qualifier to that statement.

Depending on your needs, you might prefer one behavior or the other.
Our data

- [https://osf.io/entpf/](https://osf.io/entpf/)
- Advancing Library Visibility in Africa: joint research initiative between African Library and Information Association and Institution and University of Washington
- Data about some libraries in Africa, collected in different countries
- Most of them are not on Wikidata
country

Sweden

start time

1998

1 reference
Your own tools and workflows