OpenRefine

What is OpenRefine?

"A powerful tool for working with messy data"

OpenRefine helps you to:

✓ clean data

✓ transform cleaned data from one format into another

✓ extend data with web services and external data *(for looking up related data on the internet – see video 3 on openrefine.org)*
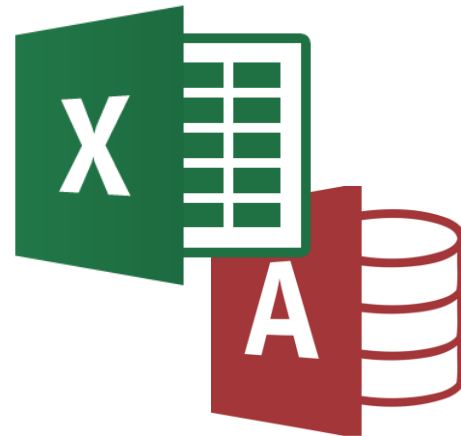
What it is not?

Database

Excel-like experience

BID GBIF

Feature comparison for data cleaning

Database

Excel

OpenRefine

Database
- Infrastructure and programing skills to edit
- Absence of easy visualization

Excel
- Usually one cell edition
- Useful to document data and perform operations
- Deficient visualization

OpenRefine
- Multiple cell edition
- Easy exploration and transformation
- Interactive visualization

Feature comparison

OpenRefine

- Multiple cell edition
- Easy exploration and transformation
- Interactive visualization

- Free
- Desktop app (offline)
- Infinite undo function
- Several type files to export/import

- Faceting / Filters
- Use of regular expressions
- Extensions
- Large community of developers (extensions, tutorials, etc)

BID GBIF

The next activity is a practical exercise using OpenRefine

file: OpenRefine_Exercise_2019.pdf