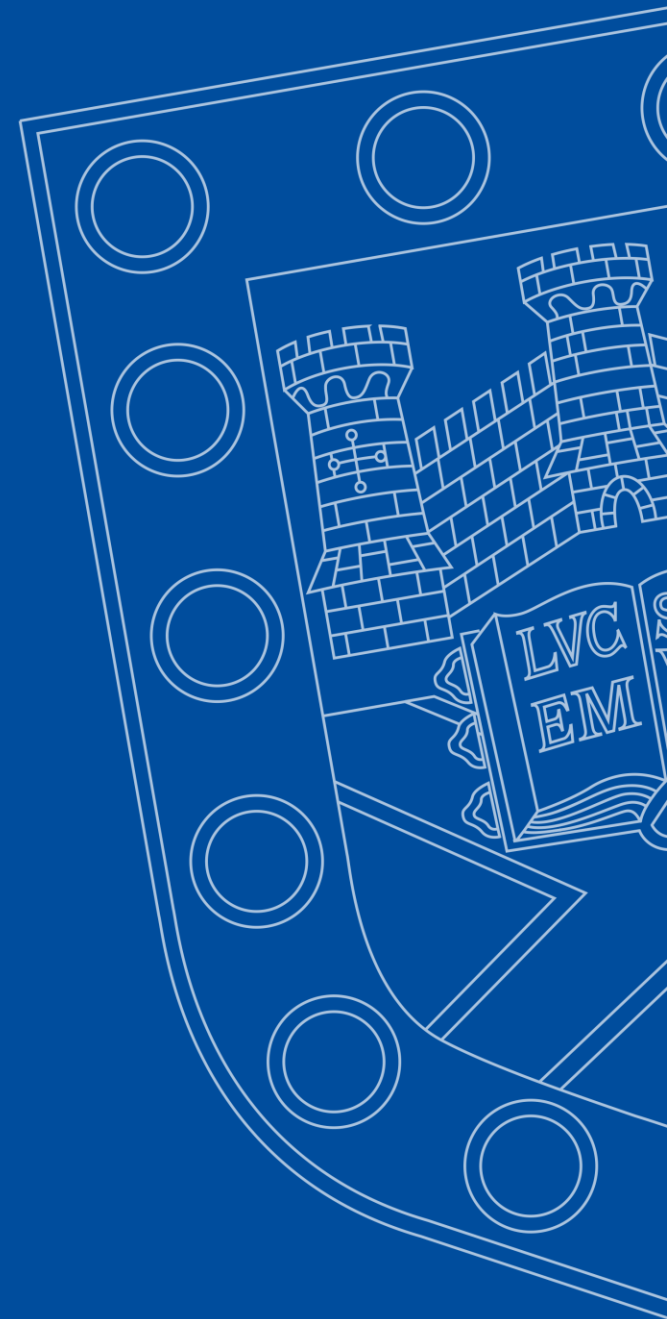




Installing OpenRefine

Exeter Digital Humanities Lab



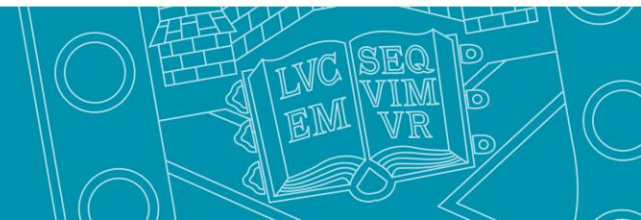
Downloading and Installing OpenRefine

Procedure

- Download from:
<https://openrefine.org/download.html>
- Ensure you get at least v 3.5.2
- On Windows, use the Windows Kit if you have installed Java from your institution or IT department
- Otherwise use embedded Java version
- *It's important to keep OpenRefine and Java updated regularly*

Considerations

- Note that OpenRefine runs as a webserver on your computer, but no internet connection is needed
- No data leaves your computer, so it's good for confidentiality
- If you're handling particularly sensitive data, make sure your local firewall blocks the relevant port (3333)



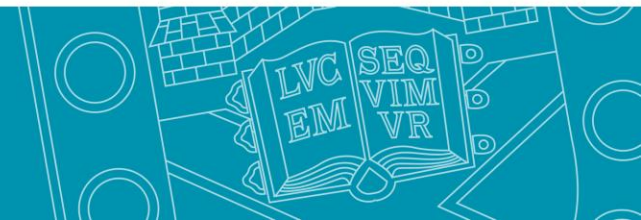
Installing OpenRefine

Procedure

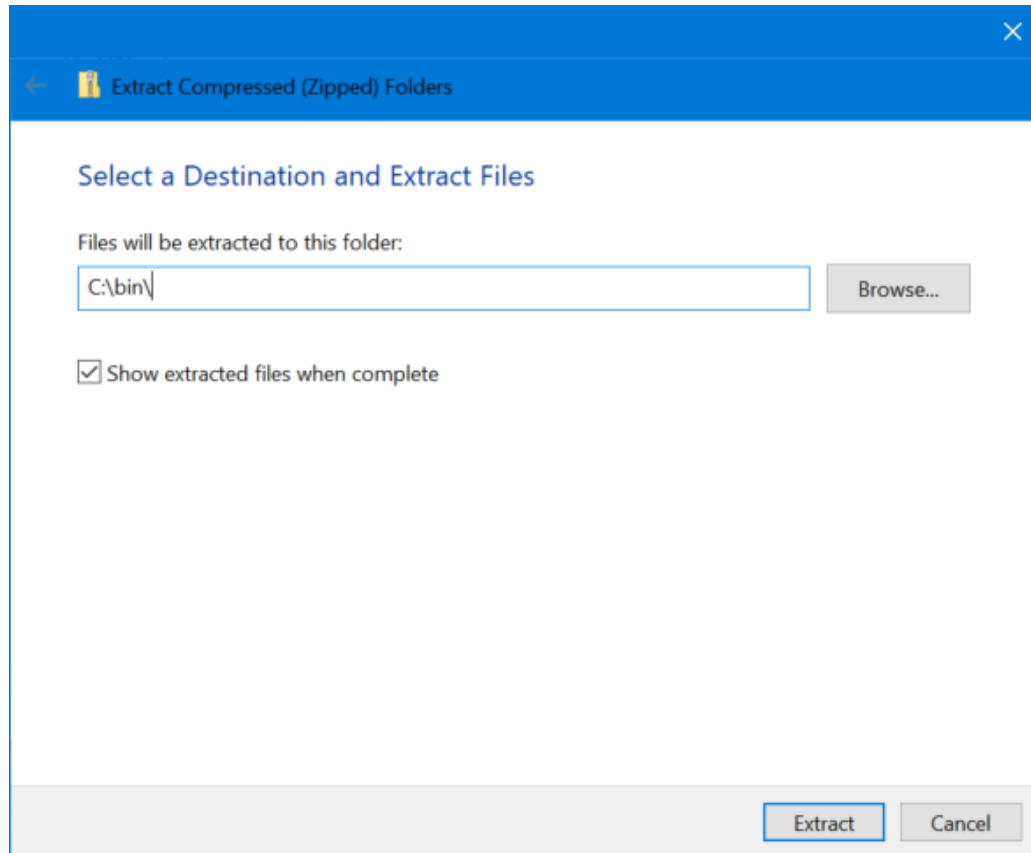
- Download the relevant kit for your operating system - this will be a .zip file
- Unpack it to a suitable location
 - If on your hard drive, will be faster, but work needs to be backed up
 - If in the cloud (e.g. OneDrive), may be slower but more secure
- Run using the instructions on the download page

Considerations

- The download is 170MB, so ensure you have space to download and install it
- You might want to create a shortcut or link to the startup script (hint, try pinning it to your taskbar)
- I tend to store OpenRefine in c:\data (windows)

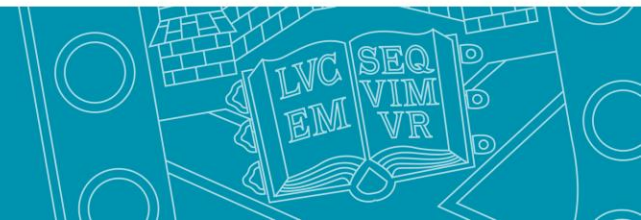


Installing OpenRefine (2)

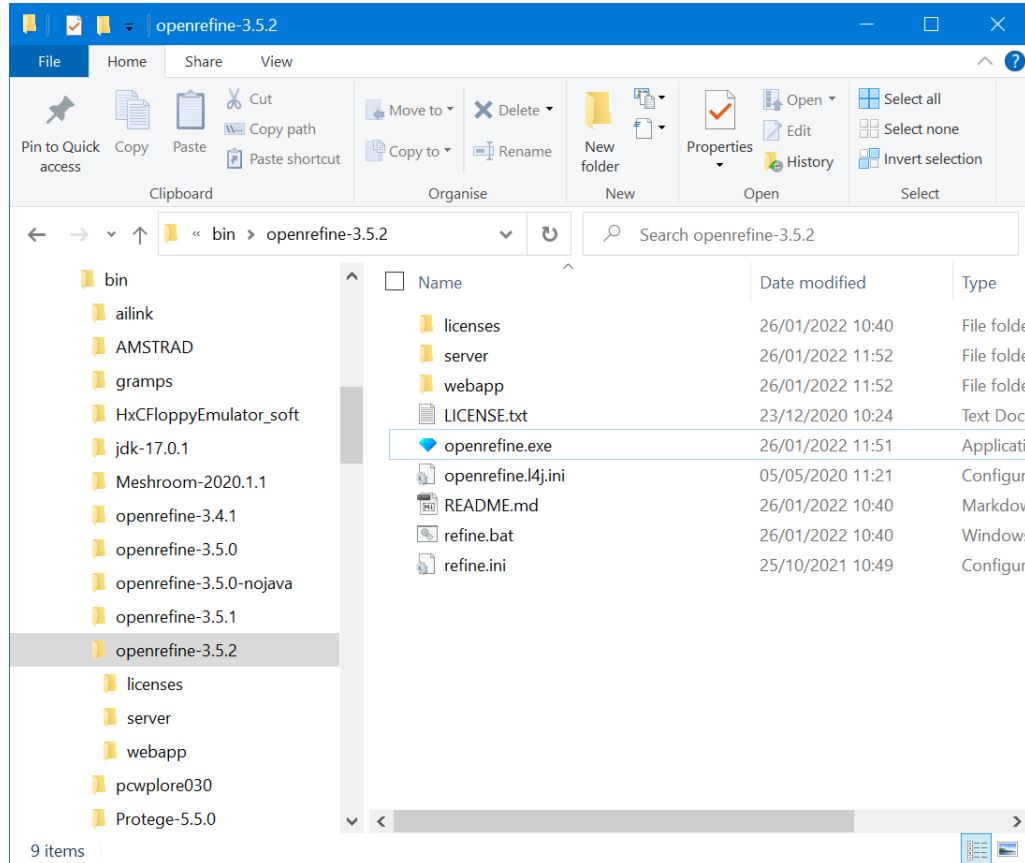


Procedure

- In Windows, right click on the downloaded file and choose 'Extract All...'
- Choose the location and extract
- The archive will unpack with a version number
- Installing new versions will automatically preserve projects
- For Macs, check the install notes



Running the Program

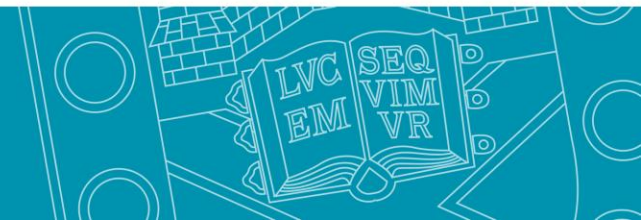


Finding the executable

- Try the OpenRefine application first
- If that doesn't work, try the 'refine.bat'
- You may get a warning that "Windows protected your PC"
 - It's safe to click on 'More info' then 'Run anyway'
 - You get this message because OpenRefine isn't 'signed' by Microsoft, but that's OK!

Working with OpenRefine

- If successful, you'll see a black text window
 - any errors will appear there
- And your default browser should open at the OpenRefine interface
 - Note that OpenRefine doesn't work **quite** as well with MS Edge than Chrome or Firefox – you can copy and paste the URL into either of these.



Working with a Dataset

- To represent a common and typical problem that OpenRefine can assist with, we are making some catalogue data available
- This is a direct Excel dump from the CALM catalogue system, but you do not need Excel or CALM to work with it
- Download sample data from here:
- <https://github.com/ExeterDigitalHumanities/openrefine/>

