Trove as a platform for digital research and creativity

University of Canberra, 3 June 2019

Before you start

1. Make sure you have a Trove API key. Once you have it, keep the tab with your key open in your browser so you can copy it when necessary.

Exploring Trove newspapers over time

In this exercise we're going to look at searches in Trove's digitised newspapers, using data from the Trove API to chart the frequency of results over time.

While you're exploring, think about the sorts of questions that you could use QueryPic to explore. What about its limitations? What sorts of things might affect the accuracy of the results?

1. Open the QueryPic Deconstructed notebook. This will open in a customised computing environment using the Binder service.
2. Note that the notebook opens in 'app mode' with all the code cells hidden. To switch into a normal notebook view, click on the 'Edit App' button. Try switching back and forward between modes to explore the difference.
3. Paste your Trove API into the box where indicated.
4. Decide on two (or more) search terms you'd like to compare, and add them, one at a time, using the query box and button.
5. **Create your chart!**
6. Once your chart has loaded, use the dropdown widget to switch the view from 'Proportion of total article' to 'Raw number of results'. What changes and why?
7. Now start to experiment with different searches and date ranges. Try comparing the same search across different states by selecting the 'Compare states' tab.
8. Once you have a chart you like, use the 'Save chart' button to save it as an HTML page. Open up the page and use your browser's 'Save page as..' option to download it to your computer.

Digging deeper with the Trove Newspaper Harvester

In this exercise we're going to create a dataset of Trove newspaper articles that will include both the article metadata and the full text. This will enable to explore the results in more depth than is possible through the web interface or the API alone.

Think again about searches you used with QueryPic. Were there any features or anomalies that could be interesting to explore? For the sake of the exercise today, we're going to limit the results returned, but you can use the Trove Newspaper Harvester to download many thousands of articles.

1. Open the Trove Harvester web app notebook, as before it will load using Binder and display in 'app mode'.
2. Paste your Trove API into the box where indicated.
3. Now go to the Trove web interface to construct a search based on your QueryPic explorations. So that we don't spend the whole workshop waiting for the harvest to complete, use the facets to limit your search by date, state, or newspaper until you have under 1,000 results.
4. Once you've constructed your search in the web interface, simply copy the url, and paste it where indicated in the Harvester notebook.
5. Check the box that says 'Save full text'. This will save the full text (OCRd) content of each article into a separate text file.
6. **Start your harvest!**
7. Once your harvest is complete you'll see a link to download the results. Download the zip file to your computer.
8. At the bottom of the harvester you'll see a link to open another notebook, click to open it. If the notebook opens in 'app mode' click the 'Edit App' button to display all the cells.
9. Work your way down the notebook by hitting **Shift+Enter** to run each cell. Can you find anything interesting? What sorts of questions might you be able to ask using these sorts of tools?

Analysing texts with Voyant

The exploratory notebook showed some ways in which you could analyse the titles of newspaper articles. In this exercise we're going to use Voyant, a web-based text analysis tool, to explore the full text of the articles that we harvested above.

1. Find the zip file you downloaded above and unzip it if necessary. Inside the zipped folder you'll find a folder labelled 'texts' that contains all the individual text files from the harvested newspaper articles.
2. Go to the Voyant instance hosted by tinker.edu.au (if that doesn't work, try the original Voyant).
3. Click on the 'Upload' button then select all the text files you harvested.
4. Voyant will open in 'Corpus' view, which combines a number of frequently used tools.
5. The 'Cirrus' tool provides a familiar word cloud, but try out 'Terms berry', or 'Context'. What do they tell you? Note that because the names of the text files all start with the date, Voyant loads them in chronological order, so tools like 'Terms' offer another way of seeing change over time.
6. Click on the window icon that appears when you hover over the top menu bar to see the full range of tools and visualisations.
7. It's easy to share links to your visualisations – just click on the link icon that appears when you hover over the menu bar of any tool.