

TLT Voyant Directions

Best practices

You are accessing pre-loaded corpora in this workshop, but there are some best practices you should follow to make sure your text is ready to work with in the future.

1. Voyant works most efficiently with uploaded text as opposed to direct URLs
2. The smaller your files, the quicker they will upload. Since you can only analyze text with Voyant, it may be most efficient to work with .txt files.
3. Think about what sorts of information you want to glean from your text. For example, are you comparing multiple versions of the same text from different publication years? If so, name your files with the date at the beginning of the filename so you can easily tell which document is which.
4. If you are using text downloaded from a website, or from a repository (like the Internet Archive), it may have content you don't want (such as html tags, or info about the library that digitized the original book). You may need to clean up your text before uploading to Voyant because you can't edit it once it's uploaded unless you replace the document.

Setting up

1. Go to the Research Guide at https://researchguides.loyno.edu/text_workshop
2. Go to "Data Sources," "1970s Earth, Wind & Fire Lyrics" to open the corpus in Voyant. You can also go to <https://voyant-tools.org/> and upload your own corpus.

Understanding your corpus skin

Earth, Wind & Fire lyrics from their 1970s albums were copied and pasted from AZLyrics (<https://www.azlyrics.com/e/earthwindandfire.html>) into individual text files. Each text file was named according to the following file structure yyyy_songTitle. Go to the folder "earthwindandfire1970slyrics" in <https://tinyurl.com/loynotext> to see the folder and file structure (you will need to sign into your Loyno Gmail to access).

At the top right of your dashboard there are five buttons:

- Language Interface Options
- Home (to start a new corpus)
- Export (to generate a stable URL for your entire corpus, an individual tool, or an image file of a visualization)
- Tool (to change to a different tool view)
- Help

Your corpus dashboard comes pre-loaded with the following tools:

- Cirrus: a kind of word cloud showing the most frequent terms
- Reader: an efficient corpus reader that fetches segments of text as you scroll
- Trends: a distribution graph showing terms across the corpus (or terms within a document)
- Summary: a tool that provides a simple, textual overview of the current corpus
- Contexts: a concordance that shows each occurrence of a keyword with a bit of surrounding context

Spend a few minutes looking at the different tools currently available in your corpus.

Modify and Download

Did you accidentally upload a wrong document, or see something you need to edit in a document? You can add, remove, and reorder any documents in your corpus. You can also download your entire corpus.

1. In the Summary tool (bottom left), click Documents
2. Click the Modify button to add, remove, or reorder documents. Note that this will create a new corpus.
3. Click Download to download a copy of your corpus locally (this will just be your source documents, not the Voyant tools and views). Useful if you're working with a pre-loaded corpus.

Stop words

The first tool panel on your dashboard should be the Cirrus, or word cloud. Do you see any words you don't want included? These are called "stop words" and can be filtered out of your results.

1. Hover your cursor over the top of the Cirrus tool. Click the button to "Define options for this tool."
2. Under "Stop Words," select "English."
3. Click "Edit list" to add additional stop words. For example, these song lyrics include a designation for the "Chorus" that we don't want included in our analysis. Add "chorus" and any other stop words to the bottom of the list, then click Save
4. Click "edit globally" to apply these stop words to all of the tools
5. Click confirm

Scale

You can toggle back and forth between looking at your entire corpus, or looking at individual documents.

1. At the bottom of a tool (like the Cirrus tool), click the drop-down menu that says "Scale." You can either choose "Corpus" for all documents, or select individual documents.

Tool Interactions

An essential part of Voyant is that events in one tool can cause changes in other tools (the exact interactions depend on a number of factors, including which tools are visible). For instance, try the following sequence in the window above:

1. click on a word (like "heart") in Cirrus, the tool in the upper left; notice how the Trends tool (upper right) now shows just the word you clicked
2. Click "reset" on the Trends tool; now click on one of the discs in the Trends tool
3. notice that the Contexts tool (bottom right) has updated with just the clicked word
4. click the first row of the Contexts tool (bottom right)
5. notice how the Reader tool has jumped to a location where that word appears, highlighted

Search

Voyant defaults to showing you the most used words in your corpus or document, but you can search for specific words within most tools; just look for the search box at the bottom of the tool. Hover over the Help icon for tips on using wildcards and other search syntax

New Tools

You can replace any of the initial tools with another one. To change to a different tool, hover over the top of the tool and select the “Tool” icon. Some tools also have individual options that you can change. Tool descriptions are available at <https://voyant-tools.org/docs/#!/guide/tools>. Here are a few to try:

1. Mandala: shows the relationships between terms and documents
 - a. Un-check “labels” to make your mandala less cluttered
2. Stream Graph: depicts the change of the frequency of words in a corpus (or within a single document)
 - a. Click the drop-down menu for “Scale” and select Documents-->1975_Reasons
 - b. Click the right arrow for more options, and play with the “Segments” scale
3. Phrases: shows repeating sequences of words organized by frequency of repetition or number of words in each repeated phrase
 - a. Change “Length” to 3; select “in the sky”
4. Veliza: like SIRI or Alexa for Voyant
 - a. Select “from text” to see what Veliza is prompted to say from your corpus
5. Word Tree: allows you to explore how keywords are used in different phrases in the corpus
 - a. Search for “boogie”

Questions?

Export

You can copy the URL to your corpus and share it with others, or bookmark to work on it later. You can also export an HTML snippet to embed the entire corpus or individual tools in another site, or export tool views as image files.

1. To export the entire corpus: at the top of your corpus, click the Export icon, expand the Export View section, select an HTML snippet radio button, click the Export button, and copy the snippet that appears in the text box
2. To export a single tool: at the top of the tool, click the Export icon, expand the Export View section, select an HTML snippet radio button, click the Export button, and copy the snippet that appears in the text box
3. To save an image from a tool: at the top of the tool, click the Export icon, expand the Export Visualization section, select either PNG or SVG, click Export

Generally speaking, a corpus will remain accessible as long as it accessed at least once a month.

Explore on Your Own

1. Select another text source from the Data Sources box on the Research Guide at https://researchguides.loyno.edu/text_workshop
2. Spend some time on your own mining your text
3. Share:
 - a. **What’s something interesting you learned from your text?**
 - b. **What’s something you want to learn from your text but don’t know how?**
 - c. **How could your liaison area(s) use this tool?**
 - d. **Bonus! Tweet an export from Voyant with the hashtag #TweetsFrom411**