Visualizing with Text: Bringing Content to the Fore to Better Assess (Mis)information

Richard Brath[1]
[1] Uncharted Software Inc., 2 Berkeley St. #600, Toronto, ON
ORCID: 0000-0001-6006-2092

Abstract

In an age of Large Language Models, online misinformation, plagiarism, hate speech, and more—the analysis of text is ever more important. Our traditional statistical visualization tools, however, have lagged behind—geared towards the visual display of quantitative information—rather than text-centric unstructured data. Visualizing with Text characterizes the design-space for directly integrating text and visualization in each other. It builds on the traditional visualization pipeline familiar in statistics and visualization: it adds a) literal text; b) visual attributes such as font weight, x-height and many more; c) mark types ranging from individual alphanumeric characters to paragraphs; and d) representations from extended traditional visualizations such as scatterplots, line charts and treemaps with text marks; to text-centric visualization such as tables, dictionaries and mind-maps. There are many examples of text-centric visualizations as applied to qualitative data and misinformation: including a scatterplot of social media vs. mainstream media; a line chart of tweet popularity; a treemap of human rights; a chart of LLM verbatim responses; mindmaps of LLM knowledge extents; and a table of text for comparison.

Key Words: text visualization, text analytics, NLP, large language models.

1. Introduction

Historically, there has been a strong emphasis in data visualization focusing on quantitative visualization (e.g. Tufte's *Visual Display of Quantitative Visualization*, Bertin's *Semiology of Graphics*, or Wilkinson's *Grammar of Graphics*). This has been exacerbated by statements such as "What gets measured gets managed", attributed to management expert Peter Drucker.

Need for qualitative analysis. There is a strong need for qualitative data analysis, as articulated by sociologist William Cameron: "Not everything that can be counted counts, and not everything that counts can be counted." The proliferation of misinformation is difficult to quantify: qualitative analysis may be necessary to assess the nuance of the content. There are many indications of the need for qualitative analytics in business as well, for example:

- Subjective data may align poorly with quantitative data, e.g. 0.53 correlation between users' subjective rating vs. user performance in Nielsen Norman Group.
- Subjective goals are as important as quantitative goals in career choices (Groysberg and Abrahams).
- When [quantitative] data and anecdotes disagree, the anecdotes are usually right (Jeff Bezos).

Challenges of LLMs. With the introduction of generative AI such as Large Language Models (LLMs), there are more textual analysis problems arising. For example:

- *Plagiarism.* GenAI generates output based on learned input, able to reproduce text and styles of authors and artists (e.g. Noam Chomsky).

- *Hallucination*. LLMs may give erroneous answers, sometimes with significant consequence (e.g. Google AI errors cause share selloff as reported in The Guardian).

Critique of word clouds. Given that there may be qualitative data, then, can data visualization be used? The highly popular text-based visualization is the word cloud uses visualization poorly: it uses arbitrary location of words, arbitrary colors, arbitrary orientation, and the single variable represented by word size typically has no legend, that frustrating any possibility of accurate decoding by the viewer. Furthermore, it uses language processing poorly as well: by discretizing information into individual words, all context is lost. For example, in a word cloud derived from Alice in Wonderland (figure 1), Mock is separated from Turtle, repeated phrases e.g. "off with her head," are completely lost, etc.

Figure 1: A word cloud is a poor visualization: position, orientation and color are arbitrary, and much context is lost by discretizing running text into individual words.

But there is much more possible. Given the starting point of the text of Alice in Wonderland, the author has found more than 50 non-word-cloud visualizations of Alice, most of which use text throughout the visualization (e.g. Brath, Surveying Wonderland). There is a need for text-centric visualizations and approaches to create them.

2. Background: Visualizing with Text

2.1 Traditional data visualization conceptualization

Many data visualization reference works identify encoding data values into visual attributes as a core step in the creation of data visualizations (e.g. Bertin, MacKinlay, Munzner, Wilkinson). This usually includes:

a) *Data*, which may be categoric, ordered and quantitative
b) *Visual attributes*, such as size, color, position, orientation
c) *Marks*, which may be points, lines or areas
d) *Layout*, which is the composition of marks with attributes into a spatial layout such as a Cartesian grid, a network, a space-filling representation, map, etc.

There are many problems with this conceptualization as applied to text, for example, prose text must be converted into the above data types. Hence Alice in Wonderland gets converted into discrete words (i.e. catetgories which are counted). Or, text does not fit within visual attributes or marks – typically text is applied after the visualization is created, such as an annotation or label.

One objection may claim that text is not visualization. Visual attributes, such as size and color, can be perceived preattentively (which is fast and automatic), whereas text requires reading (which requires directed attention and is slow). Figure 2 shows on the left some simple plots using preattentive visual attributes, and on the right some textual plots, varying in text and in typographic attributes such as font weight and case.

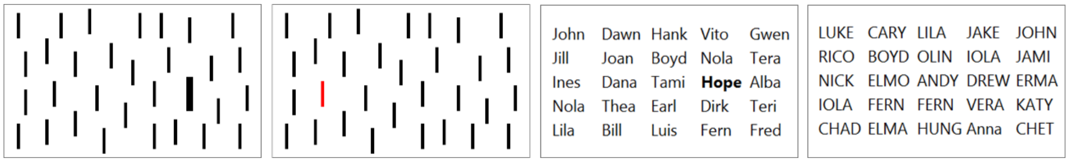


Figure 2: One item is different that the others. Left: difference using preattentive visual attributes size and color. Right difference in typographic attributes of font weight and case.

Note that the differences are perceivable without reading the text – and this technique has been used on maps for hundreds of years. Furthermore, reading can be automatic as shown by the Stroop effect (1935). Automaticity is the ability to perform a well-practiced task with low attentional requirements (Bargh 1994).

2.2 Some non-traditional visualizations use text extensively

As noted above, maps use a lot of text and the text encodes data using typographic attributes such as font weight and capitalization. A review of historical and non-statistical visualizations shows a wide variety of text utilization within visualizations encoding information, such as genealogical diagrams, tables of contents, code editors, artworks (e.g. Cheng), and mindmaps (see Brath, Visualizing with Text, for more examples). A few examples are shown in Figure 3.

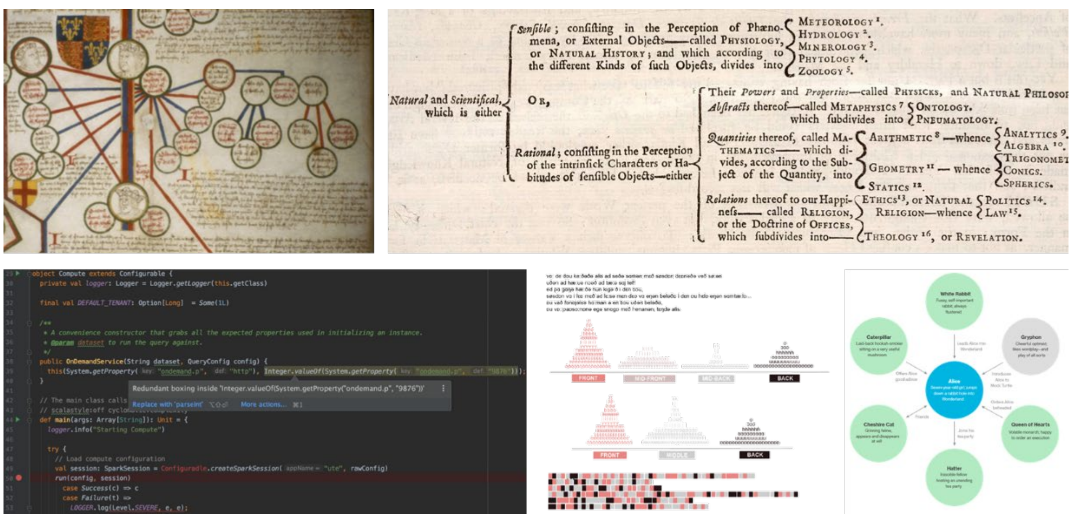


Figure 3: Interesting uses of text in visualizations include (clockwise from top left): genealogical diagrams (text color, text in nodes, intertwined commentary), table of contents (typographic formats, tree layout, phrases), code editor (typographic formats), letter distributions (individual glyphs), and a mind map (sentences in a color-coded graph).

2.3 A design space for text within visualization

Based on a review of these examples (and many others), the range of encodings significantly broaden:

a) *Data*, extended to include literal text, such as full passages
b) *Visual attributes*, extended with typographic features, such as glyphs, weight, italics, capitalization, typeface, width and so forth.
c) *Marks*, which may extend from individual characters, through words and phrases, to sentences and paragraphs.
d) *Layout*, wherein existing layouts can be extended to use textual marks with typographic or other visual attributes, and other layouts such as tables or mindmaps may be used.

There are many examples of interactive visualizations that can be created using this design space. Some examples are shown in Figure 4, from left to right: a line chart with text along each line; the text from *Alice in Wonderland* with words formatted to facilitate skimming; baseball players by team, position, and hitting stats in font color, font weight, font width; and thousands of coroner reports by cause of death. These examples are from the *Visualizing with Text* companion website.

Figure 4: Some snapshots of interactive text visualizations: microtext for lines; prose weighted to aid skimming; a table of baseball players with each name encoding multiple metrics; thousands of causes of deaths from coroners' reports.

3. Visualizing with text examples (with misinformation)

The design space is very large, with thousands of possible permutations. These few examples should aid in understanding the range of extents feasible, as well as possible use towards analysis of misinformation.

3.1 Literal labels, with diverging sentiment example

A simple example uses textual labels to encode data. The scatterplot in Figure 5 shows a scatterplot of 100 public companies' sentiment. Along the x-axis is the sentiment in mainstream media and along the y-axis is the sentiment in social media. One might expect mainstream media and social media to align – i.e. all the companies would be along the diagonal from bottom left to top right. However, there are many companies with mismatched sentiment. This mismatch could be due to misinformation or other information. For example, Volkswagen is near the bottom right – i.e. highly negative social sentiment, but somewhat positive mainstream sentiment. In this case, the divergence is due to timing: this snapshot is somewhat after *Dieselgate,* wherein Volkswagen had been caught manipulating emissions data. For social media, consumers are highly negative, coming to the realization that their Volkswagen diesel vehicles have lost significant resale value. For mainstream media, *Dieselgate* is already old news and is now reporting on newer events.

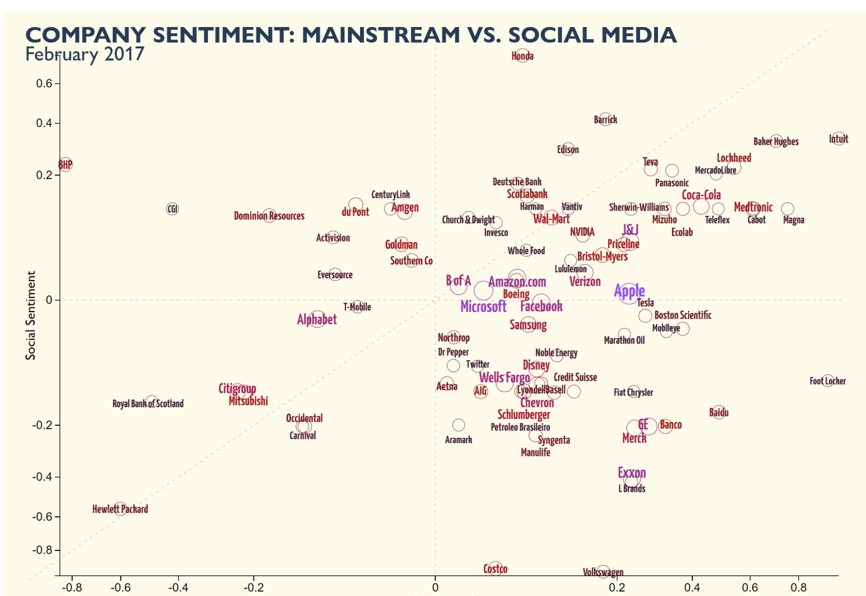


Figure 5: Company sentiment in mainstream media and social media. A divergence may indicate misinformation or other differences.

This visualization would have indicated the sentiment mismatch during the GameStop short squeeze in January 2021, with social media on reddit highly positive and mainstream media informed by shortsellers as highly negative. More broadly, this approach can be used to evaluate a mismatch between any two textual sources, For example, the interactive scatterplots in *What have language models learned* (Pearce, 2021) similarly use labeled scatterplots, to reveal biases in LLMs by gender, region and so on.

3.2 Microtext lines, with social media popularity example

Text in a visualization does not need to be limited to simple labels. The lines in the line chart in Figure 6 have been replaced with microtext tweet content. The overall line indicates the number of retweets over time showing how fast the particular tweet becomes viral. The text is immediately available for inspection. This example is top 5 tweets regarding hurricanes and the fastest growing tweet by a large margin has comedic content. Unfortunately, in an unfolding natural disaster comedy is more viral than safety.

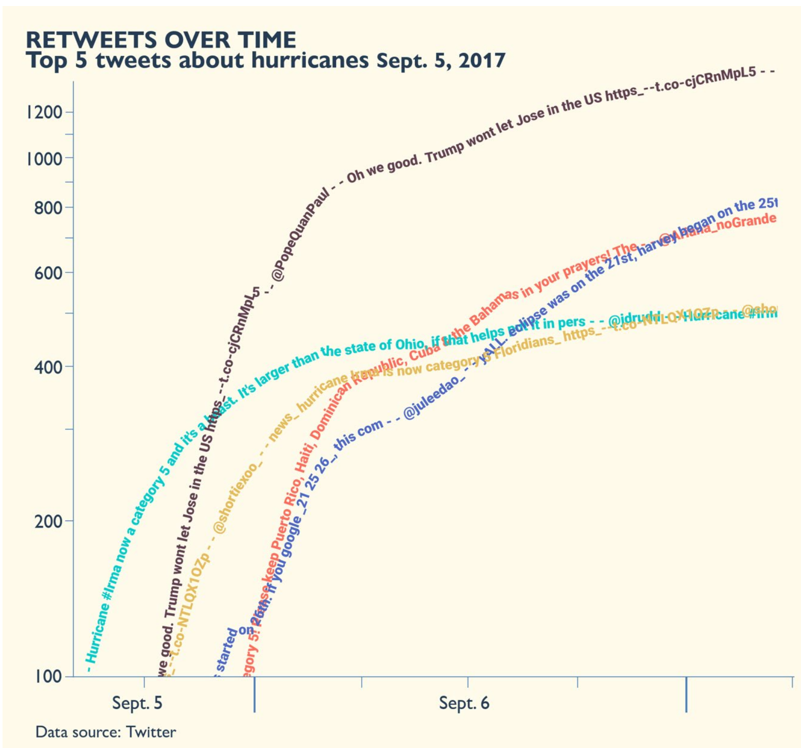


Figure 6: Company sentiment in mainstream media and social media. A divergence may indicate misinformation or other differences.

3.3 Areas with text, with example indicating relation between oil and human rights

Quantitative visualizations require the non-quantitive to be quantified to be represented. This severely limits visualizations' ability to express complex nuanced information. Consider the visualization in Figure 7. The treemap is sized to indicate a countries' oil exports. Oil exports are used to fund state initiatives, including military and/or human rights abuses. How can human rights be visualized? One can search for a singular metric, such as the *Global Peace Index*, here encoded as color. But indexes of qualitative data can be difficult to understand—they attempt to rollup a variety of disparate data. Instead, paragraphs of text can be added to the areas to complement the quantitative values. In this case, text from the lede paragraph of *Human Rights Watch* per country is displayed. Reading the text characterizes what kinds of human rights issues exist in each country. While Russia has a terrible record (most repressive regime since the Soviet era), the UAE also has issues (dissidents detained after completing sentences), as does Saudi Arabia (accountability for the murder of Jamal Khashoggi), and other poor records for big exporters. The differences are much more easily understood as direct text as opposed to an abstract numeric value.

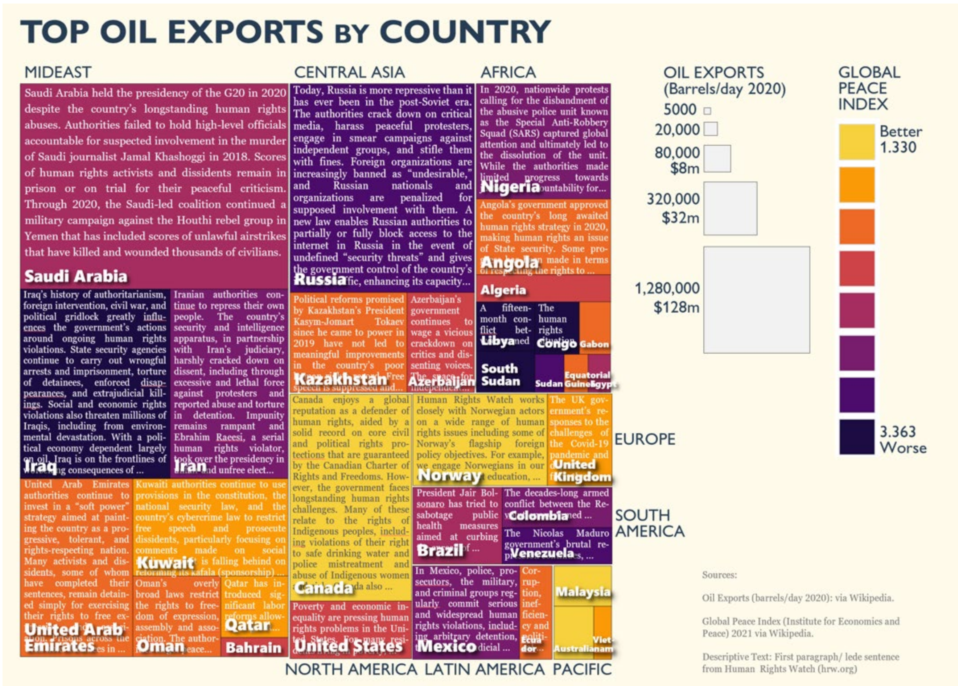


Figure 7: Treemap indicating oil exports by country, color indicates the *Global Peace Index*, and prose text indicates human rights abuses in the lede paragraph from *Human Rights Watch*.

This approach of fitting diverse text into organized areas can be an effective way of arranging a lot of qualitative feedback, such as the small portion of the poster *Landschapstekening Kunsten*, in Figure 8, (Huyghebaert and Boiron).

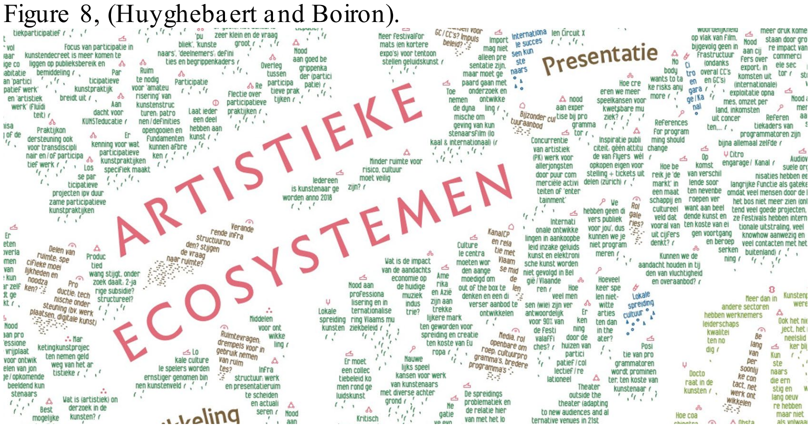


Figure 8: A hierarchical organic treemap of text regarding many trends in the arts.

3.4 Quantitative text, indicating potential LLM plagiarism

As indicated in the Introduction, one criticism of LLMs is the potential to plagiarize. This can be difficult to assess, as LLMs are non-deterministic and each answer can be different. In Figure 9, (Brath et al, 2023), an LLM has been prompted to extend a passage from *Alice's Adventures in Wonderland*: "Would you tell me, please, which way to I ought to go from here?" This prompt was asked 60 times, and each response then marked up to indicate the correct words (in green) and incorrect words (in black), then sorted by number of correct words. The top two answers are more than 100 words correct, and only one answer at the bottom has no correct words. The answers in the middle represent the "average" answer, i.e. the more probable answers—there are a number of answers of approximately 30 words and another number of answers of approximately 60 words. Thus, the visualization helps to characterize what the LLM has learned strongly.

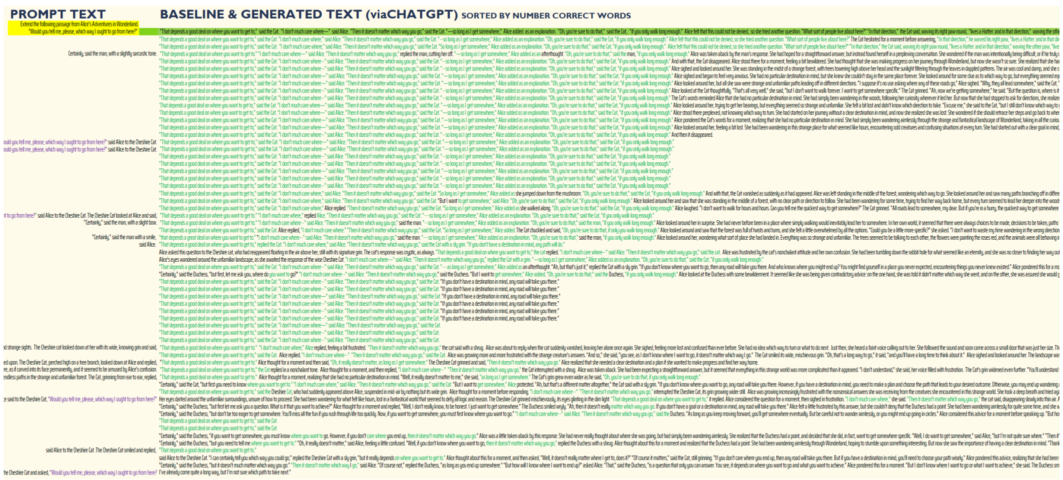


Figure 9: Each row is a response from an LLM to a prompt asking the LLM to extend a quoted passage from a popular book. The correct words in each response are in green, rows sorted by the number of correct words.

3.5 Mindmaps, showing LLM knowledge extents of the (fake) moon landing

Extending the previous approach, an LLM can prompted with a general question and then the answers analyzed. In Figure 10, the LLM has been prompted 850 times. The answers are then processed based on commonality in the responses to create a mindmap visualization. The left image shows commonality in explanations to the *first moon landing*, with hundreds of responses citing *Apollo, Niel Armstrong, Buzz Aldrin, NASA*; and, longer sentence fragments on order of 50 responses, e.g., *Apollo 11, the spaceflight that landed the first two people on the Moon*; or, *Armstrong became the first person to step on the lunar surface six hours after landing on July 21 at 02:56 UTC*; etc. Successively smaller boxes with smaller fonts are statements the LLM repeats less, such as statements about *JFK* and the *Cold War* repeated only a few times out of 7894 generated sentences—i.e. the LLM has a very low frequency of including these elements in a response and thus does not have strong learning of this information. In other words, if the LLM produces an answer about the first moon landing outside of this mindmap, it is low confidence.

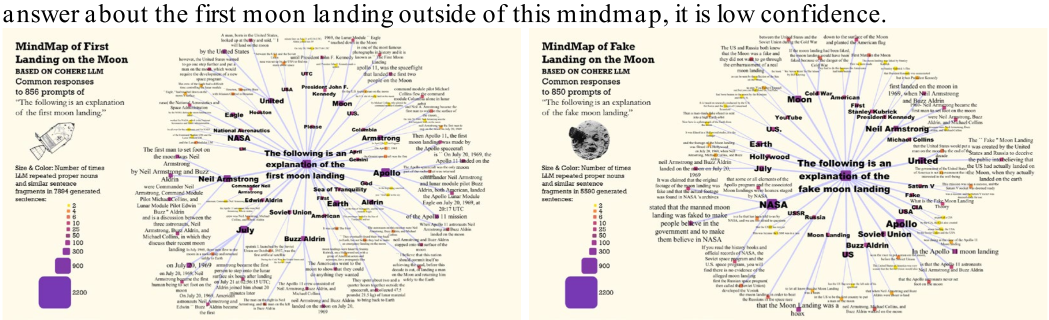


Figure 10: Mindmaps of LLM answers about the first moon landing (left) and the fake moon landing (right).

As LLMs have been trained on vast amounts of data, including the Internet, they have also been trained on misinformation contained on the Internet. The right image in Figure 9 shows a mind map of explanations to the *fake moon landing*. Entities such as *Apollo, NASA* and *Neil Armstrong* are frequent, as are *Hollywood* and the *Soviet Union*. Frequent sentence fragments include *manned moon landing was faked to make people believe in the government and NASA;* or *the US and Russia both knew the Moon was a fake and they did not want to go through the embarrassment of a real moon landing*. As such, users of LLMs need to be critical in the use of LLM output: all the errors of the Internet may be embedded in LLMs.

3.6 Comparison and multiple perspectives (via LLM)

There are many cases where text should be compared. One example is issue analysis: rarely are issues only two sided. “There is always a well-known solution to every human problem—neat, plausible, and wrong” (H.L. Mencken). Thus inspired, an LLM was prompted to create a table of disagreements on bills in U.S. Congress, including arguments for and against by the parties and a third party position, as shown in Figure 11 (color and format added by author). For example, tax cuts are criticized as a benefit for the wealthy, supported as economic stimulus, and criticized for not offsetting tax cuts with spending cuts.

SIGNIFICANT PARTY-LINE VOTES IN CONGRESS FROM 2012-2024

BILL	Democratic Position	Republican Position	Alternative Position	Senate?
American Rescue Plan Act (2021)	Supported as essential COVID-19 relief providing direct payments and unemployment benefits	Opposed as excessive spending that could fuel inflation	Libertarian Party: Opposed all COVID relief as government overreach, advocated for complete reopening and no restrictions	✓
Infrastructure Investment and Jobs Act (2021)	Backed as crucial investment in roads, bridges, broadband	Split support - Some opposed due to cost	Green Party: Criticized as insufficient for climate crisis, demanded larger renewable energy investment	✓
Inflation Reduction Act (2022)	Supported as climate action and healthcare cost reduction	Opposed as potential to worsen inflation	Progressive Caucus (within Democrats): Criticized as too modest, wanted Medicare for All and Green New Deal	✓
For the People Act (2021)	Supported to expand voting rights	Opposed as federal overreach	Constitution Party: Opposed as unconstitutional, advocated returning all election control to states	✗
John Lewis Voting Rights Act (2021)	Backed to restore VRA protections	Opposed as unnecessary intervention	ACLU (Advocacy Group): Supported but criticized as not going far enough on preclearance provisions	✗
Patient Protection and Affordable Care Act (2012)	Supported healthcare reform	Opposed as government overreach	Socialist Party USA: Criticized for not establishing single-payer healthcare system	✓
Tax Cuts and Jobs Act (2017)	Opposed as benefiting wealthy	Supported as economic stimulus	Libertarian Party: Supported tax cuts but criticized for not matching with spending cuts	✓
First Step Act (2018)	Supported reform	Mostly supported	ACLU: Supported but criticized for not addressing mandatory minimums comprehensively	✓
Build Back Better Act (2021)	Backed social programs	Opposed spending scale	Working Families Party: Supported but demanded stronger labor provisions and wealth tax	✗
American Health Care Act (2017)	Opposed ACA repeal	Supported as replacement	Tea Party Groups: Opposed as not fully repealing ACA	✗
George Floyd Justice in Policing Act (2021)	Supported police reform	Opposed undermining law enforcement	Black Lives Matter: Criticized as insufficient, called for more fundamental restructuring of policing	✗
COVID-19 Hate Crimes Act (2021)	Supported addressing hate crimes	Mixed support	Asian American Justice Groups: Supported but criticized for focusing on enforcement rather than prevention	✓

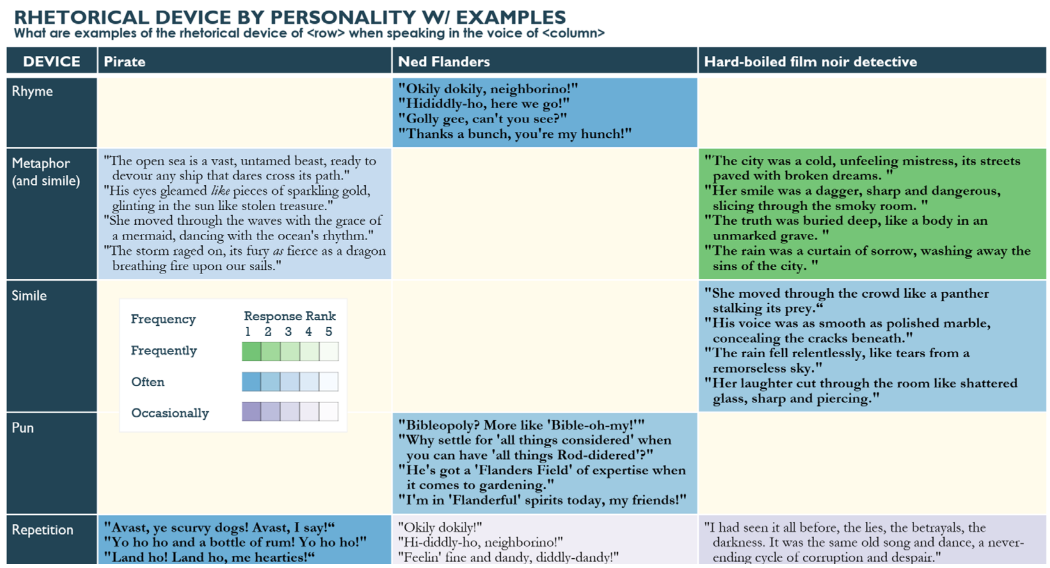
Font weight indicates difference (in %) for/against vote: thin for a thin margin, **heavyweight for big margin** 0-1, 1-2, 2-3, 3-5, **5-10**, **10-25**, **25-50**, **50-75**, **75-100**
Cell fill color indicates degree of support from aligned (unsaturated) to opposed/criticized (saturated)
Data and analysis generated by Claude AI 10/25/2024 in response to: Create a table of top disagreements by bill by votes in Congress by Republicans vs Democrats for each of the last 12 years. Present it as a table, with the bill in the first column and the strongest argument for/against in the next two columns. Also include vote counts, and an additional column, with an additional strongly expressed third party position, significantly different from the Democratic/Republican position. Include the identity of that third party if available. Formats added by author.

Figure 11: Table of bills with various arguments supporting or criticizing those bills.

The approach could be expanded in many ways, to include more positions, to count the frequency of different positions, positions of specific people, social media response, assessments by population subsets, impact forecasts, and so on.

With regards to misinformation and disinformation on the Internet, much content can hide behind fake accounts, robots, and so forth: the originator is unknown. However, authors use various stylistic cues in their text, and these cues identify their community. Earlier Natural Language Processing technology (NLP) was limited in detection of these stylistic devices: for example, rhetorical devices such as metaphors were near impossible to extract from text. LLMs can extract, use and reuse these devices, known as *text style transfer*. This same technique can be used to generate style, or, extract style from text. Figure 12 shows a few examples of frequent rhetorical devices used by various personas, with color and intensity indicating frequency and rank. For example, film noir detectives frequently use metaphors, pirates often use rhyme. Furthermore, each cell contains textual examples, which aid the viewer to further characterize the stylist device: e.g., pirates use nautical and treasure themes; the detective uses gritty urban crime references. Using these techniques on suspect text in the opposite direction can help characterize the sources of misinformation.

The comparison of many responses is also a critical task in the development of LLMs. Google’s LLM Comparator (Kahng et al, 2024) uses tables of responses to assess the quality of responses by variants of LLMs (or variants of prompts).



RHETORICAL DEVICE BY PERSONALITY W/ EXAMPLES
What are examples of the rhetorical device of <row> when speaking in the voice of <column>

DEVICE	Pirate	Ned Flanders	Hard-boiled film noir detective
Rhyme		"Okily dokily, neighborino!" "Hididdly-ho, here we go!" "Golly gee, can't you see?" "Thanks a bunch, you're my hunch!"	
Metaphor (and simile)	"The open sea is a vast, untamed beast, ready to devour any ship that dares cross its path." "His eyes gleamed *like* pieces of sparkling gold, glinting in the sun like stolen treasure." "She moved through the waves with the grace of a mermaid, dancing with the ocean's rhythm." "The storm raged on, its fury *as* fierce as a dragon breathing fire upon our sails."		"The city was a cold, unfeeling mistress, its streets paved with broken dreams. " "Her smile was a dagger, sharp and dangerous, slicing through the smoky room. " "The truth was buried deep, like a body in an unmarked grave. " "The rain was a curtain of sorrow, washing away the sins of the city. "
Simile			"She moved through the crowd like a panther stalking its prey." "His voice was as smooth as polished marble, concealing the cracks beneath." "The rain fell relentlessly, like tears from a remorseless sky." "Her laughter cut through the room like shattered glass, sharp and piercing."
Pun		"Bibleopoly? More like 'Bible-oh-my!'" "Why settle for 'all things considered' when you can have 'all things Rod-didered'?" "He's got a 'Flanders Field' of expertise when it comes to gardening." "I'm in 'Flanderful' spirits today, my friends!"	
Repetition	"Avast, ye scurvy dogs! Avast, I say!" "Yo ho ho and a bottle of rum! Yo ho ho!" "Land ho! Land ho, me hearties!"	"Okily dokily!" "Hi-diddly-ho, neighborino!" "Feelin' fine and dandy, diddly-dandy!"	"I had seen it all before, the lies, the betrayals, the darkness. It was the same old song and dance, a never-ending cycle of corruption and despair."

Figure 12: Small portion of a table indicating frequency of rhetorical device by persona with example quotes.

4. Constructing and evaluating text-centric visualizations

Text-centric visualizations are not pervasive. These techniques do not exist as popular ready-made tools. Some open-source software can handle longer text phrases: e.g., Cytoscape is graph-drawing software used, in part, to create Figure 10. The output in Figure 9 was partially processed using an Excel macro. Figures 5 and 6 use the open-source library D3.js with Javascript code—about 100 lines of code are required for each. Alternatively, LLMs are improving at generating visualizations. Figure 14 was generated with a prompt to Claude, roughly as follows: *Here is a dataset <tab delimited data>. Create a scatterplot. Use year on the x-axis, emotion on the y-axis. Include labels with each item, set the font size to frequency, and color to emotion. Use associative colors, e.g. red for anger, and include a legend*. The LLM struggled to improve the title and legend, these were added manually. However, each generation of LLMs will improve: as per Ethan Mollick in *Co-Intelligence: Living and Working with AI*, “Today’s AI is the worst AI you will ever use.”

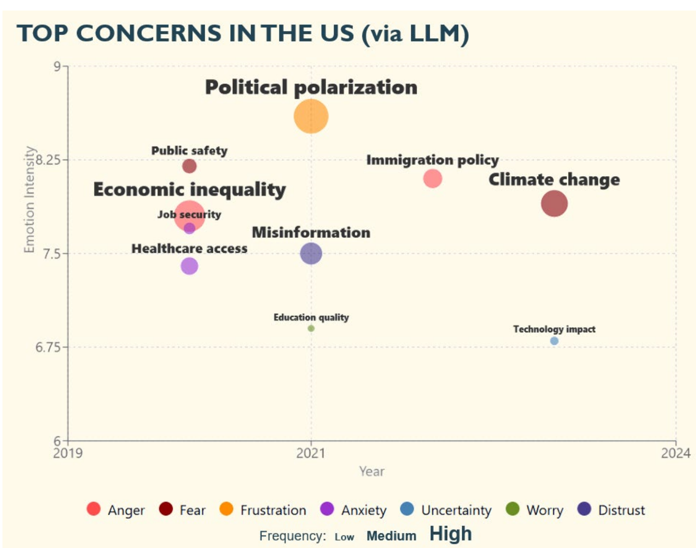


Figure 13: Table of bills with various arguments supporting or criticizing those bills, generated by an LLM, Oct 2024.

The effectiveness of text-visualizations can be assessed with various quantitative and qualitative evaluations, such as measurement of tasks, or following the nested model for visualization (user goals, user task definition, appropriate encodings, appropriate algorithms—see Munzner). Lang and Nacenta (2022) have done extensive perceptual studies evaluating typographic encodings, largely confirming the above techniques can be effective.

5. Conclusion

Visualization of qualitative data, including significant text, has always been feasible as historic examples illustrate. A framework for the use of text within visualization helps to frame the breadth of possibilities. The examples show a variety of applications for textual analyses, including examples related to information potentially subject to manipulation or errors whether by deliberate misinformation, or by faults from textual processing and generation, such as Large Language Models. While a single toolset for generating the breadth of text visualization does not exist, there are many potential approaches to creating these visualizations.

There is much future work. The examples of misinformation visualization above show how text visualizations *could* show misinformation. Stronger analytical techniques, both qualitative and quantitative, should be used together with the visualizations to create more thorough analyses. Most of the examples have some interaction, such as tooltips, filtering, zoom, and so on. Much more interaction is feasible and could aid more through analytical workflows, for example, drill-through to original sources, on-demand explanations, and so on.

References

Alice in Wonderland Character Map. Accessed 10/27/2024. coursehero.com/lit/Alice-in-Wonderland/character-map/

Bargh, John A. *The Four Horsemen of Automaticity: Awareness, Intention, Efficiency, and Control in Social Cognition*. In Handbook of Social Cognition. R. Wyer and T. Srull eds. 1994. 1—40.

Bertin, Jacques. *Semiology of Graphics*. University of Wisconsin Press, 1983.

Bezos, Jeff. *Jeff Bezos: Amazon and Blue Origin*. On Lex Fridman Podcast #405. lexfridman.com/jeff-bezos-transcript/#chapter7_principles

Brath, Richard. *Visualizing with Text*. CRC Press, 2021; and companion website: richardbrath.wordpress.com/books-and-chapters-by-richard-brath/visualizing-with-text-book-companion-web-site/

Brath, Richard. *Surveying Wonderland for many more literature visualization techniques.* At IEEE VisWeek 2021, 6th Workshop on Visualization for the Digital Humanities (VIS4DH), October 2021. arXiv preprint arXiv:2110.08584

Brath, Richard, Adam Bradley, and David Jonker. *Visualizing Textual Distributions of Repeated LLM Responses to Characterize LLM Knowledge.* At NLViz Workshop 2023 at IEEE Vis2023, Melbourne, Australia. 2023. uncharted.software/assets/visualizing-textual-distributions-of-repeated-LLM-responses.pdf

Brath, Richard, Adam Bradley, and David Jonker. *Visualizing LLM Text Style Transfer.* At NLViz Workshop 2023 at IEEE Vis2023, Melbourne, Australia. 2023. https://uncharted.software/assets/visualizing-LLM-text-style-transfer.pdf

Cameron, William Bruce. *Informal sociology: A casual introduction to sociological thinking.* Random House, 1963.

Chambers, Ephraim. *Cyclopaedia, or, An Universal Dictionary of Arts and Sciences*, 4th ed. 1741.

Cheng, Yi-Chia. *Down the Rabbit Hole: Visualizing Linguistic Distance And Relationships With Alice in Wonderland*, MFA thesis, North-eastern University, Nov 2019. repository.library.northeastern.edu/files/neu:m0455c25q

Chomsky, Noam. *Noam Chomsky Says ChatGPT is a Form of High-Tech Plagiarism*, NYTimes, Feb 17, 2023.

Drucker, Peter. As quoted in *Top 10 Peter Drucker quotes for business leaders.* CIO Views. Accessed 10/26/2024. cioviews.com/top-10-peter-drucker-quotes-for-business-leaders

Genealogy of the kings of England to Richard III. Chronicle of the Percy family. MS Bodl. Rolls 5, Oxford, Bodleian Library, c. 1485. digital.bodleian.ox.ac.uk/objects/f889e628-81b0-4201-b86c-b8ece4c1f65c/

Groysberg, Boris and Robin Abrahams. *What Does Success Mean to You?* Harvard Business Review, 2014. HBR-GA14. hbr.org/2014/02/what-does-success-mean-to-you

Huyghebaert, Pierre and Sophie Boiron. Landschapstekening Kunsten, Speculoos 2019. speculoos.com/fr/mapdesign/bouwmeester-brussels.html and kunsten.be/nu-in-de-kunsten/introductie-over-de-landschapstekening/

Kahng, Minsuk, Ian Tenney, Mahima Pushkarna, Michael Xieyang Liu, James Wexler, Emily Reif, Krystal Kallarackal, Minsuk Chang, Michael Terry, and Lucas Dixon. *LLM Comparator: Visual Analytics for Side-by-Side Evaluation of Large Language Models*. ACM, 2024. arxiv.org/abs/2402.10524

Lang, Johannes, and Miguel A. Nacenta. *Perception of letter glyph parameters for infotypography.* ACM Transactions on Graphics (TOG). 41.4 (2022): 1-21.

Mackinlay, Jock. *Automating the design of graphical presentations of relational information*. ACM Transactions On Graphics (Tog) 5.2 (1986): 110-141.

Munzner, Tamara. *Visualization analysis and design*. CRC press, 2014.

Nielsen, Jakob. *User Satisfaction vs. Performance Metrics*. Nielsen Norman Group. 2012. nngroup.com/articles/satisfaction-vs-performance-metrics/

Pearce, Adam. *What Have Language Models Learned?* Pair Explorables, Google Research, 2021. pair.withgoogle.com/explorables/fill-in-the-blank/

Stroop, John Ridley. *Studies of interference in serial verbal reactions*. Journal of Experimental Psychology, 18(6), 1935. 643–662.

The Guardian. *Google AI chatbot Bard sends shares plummeting after it gives wrong answer*, The Guardian, Feb 9, 2023. www.theguardian.com/technology/2023/feb/09/google-ai-chatbot-bard-error-sends-shares-plummeting-in-battle-with-microsoft

Tufte, Edward. *The Visual Display of Quantitative Information.* Graphics Press, 1983.