Article Review

- **Sai Priya Jyothula**

Paper selected: Kim, Hyoyoung, and Jin Wan Park. "Text Visualization: Expressive Materials and Diverse Approaches." In 2013 17th International Conference on Information Visualisation, pp. 74-79. IEEE, 2013

Context:
This paper aims to introduce a different perspective of text visualization – text visualization not just as a means of information delivery but also as a whole new genre of art. The authors include various examples of text visualization to demonstrate how the method of expression varies in text visualization.
The audience of this paper are people that are new to visualization as well as the existing visualization community, data scientists, academics in science, art, liberal arts and sociology.

Contributions:
The paper discusses different purposes of text visualization – to convey information about text form, structure and content, to derive patterns, relationships and trends of text data, for creation of storytelling and as a form of artistic expression. The authors also include diverse examples to demonstrate how text visualization serves these purposes. Basically, the paper focuses on providing the reader with a broad sense of the tools and approaches used in text visualization.

Structure of the paper:
The authors introduce why text visualization is extensively used and then talk about text data as a tool for text visualization and discuss its types and characteristics. The paper then goes on to discuss about the various methods of expression depending upon the purpose or intention of visualization. Different examples demonstrating the various methods of expression used for the same data from the same text are mentioned. Later on, the paper explains different approaches to text visualization with the help of existing examples and finally concludes saying that more systematic research is required in the field of text visualization.

Results:
In the paper conclusion, the need for systematic examination and theoretical analysis of text data and text visualization is justified and emphasized. The paper doesn't provide any quantitative results but introduces some interesting research ideas to think about.

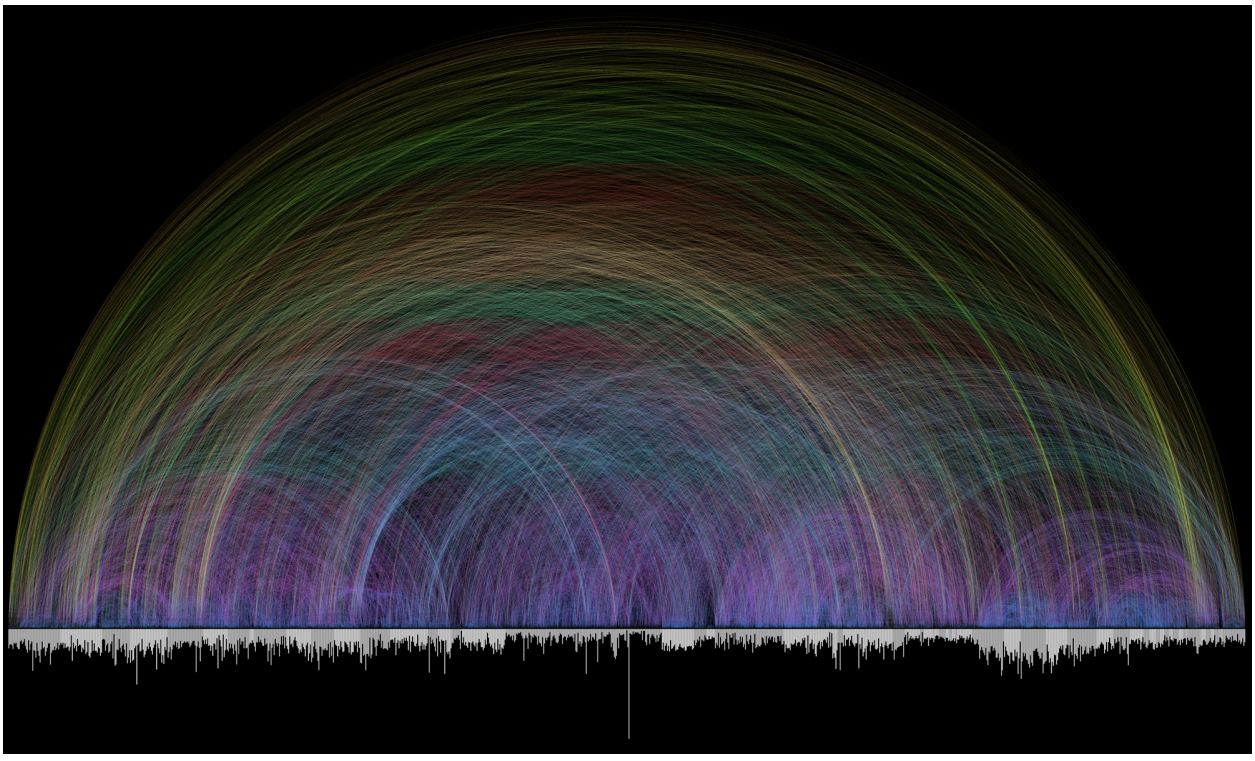
Methods used to derive results in the paper:
The paper basically provides an overview of text visualization by mentioning various example visualizations that match with the concepts discussed in the paper. For example, they included a TextArc to demonstrate the visualizations that are used to

deliver information regarding the content of the text. Though the examples are incredibly helpful to understand the different approaches and methods of expression mentioned in the paper, the lists of the different approaches used and methods of expression are by no means exhaustive or comprehensive.

Technologies:
The paper does not contain references to any new or innovative technologies that the authors used to visualize text data. On the other hand, the authors mention many other notable existing techniques used by the visualization community, such as TextArcs and word banks (based on literary organism structure by Jack Kerouac) to name a few.

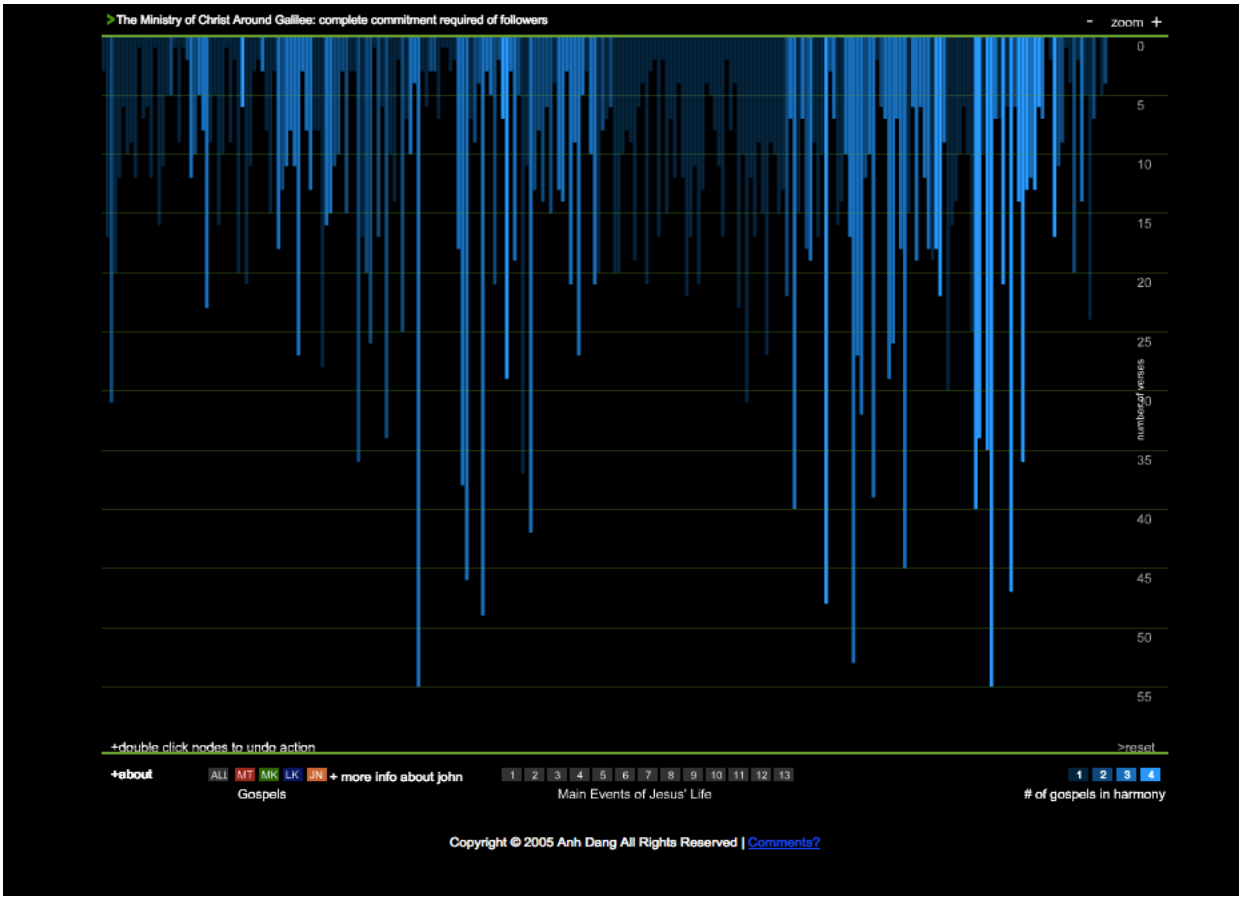
Figures:
The authors do not take the help of any figures or visualizations to explain the ideas in the paper. But the paper contains references to existing visualizations and here's my explanation for each one of them:

Visualization 1: Visualizing Bible Cross-References by Chris Harrison and Christoph Römhild



In this visualization, Bible text is used. The bar graph that runs along the bottom represents all of the chapters in the Bible. Chapters alternate in color between white and light grey. The length of each bar denotes the number of verses in the chapter. Each one of the 63,779 cross-references found in the Bible is depicted by a single arc—the color corresponds to the distance between the two chapters.

Visualization 2: Gospel Spectrum by Anh Dang



The text data used is Bible text. The bars on the graph vary in length, brightness, and color based on the data they represent. All bars are placed in a chronological order as they appear in the books and contain information such as scripture references, number of verses, number of gospels in harmony, and full text.
Length: Total number of verses describing an event.
Brightness: Total number of gospels describing an event. The brighter the bar, the more authors mention that particular event.
Colors: The colors representing each book, Matthew, Mark, Luke, and John are Red, Green, Blue and Orange respectively.

Visualization 3: New Testament Social Network (IBM many eyes)

Visualization 4: Textual Visualization based on Readability by Hyoyoung Kim, Jin Wan Park

Visualizations 5, 12 and 13: Visualizations by Stefanie Posavec

Visualization 6: TextArc by W. Bradford Paley

Visualization 7: Visualizing the Text of Phillip Pullman's Trilogy "His Dark Materials" by Tim Legan, Linda Baker

Visualization 8: Map of Science by Martin Rosvall, Carl Bergstrom

Visualization 9: Email Map by Christopher Paul Baker

Visualization 10: 7 Days of Source Day #2: NYTIMES 365/360 by Jer Thorp

Visualization 11: Visual Genealogy by Jin Wan Park

Confusion or ambiguity in the paper:
In the introduction of the paper, the authors talk about the extensive research in the area of text visualization but no references to any links or articles are made. It would have been helpful if there are some references so that the context of the paper would have been easily understood and similar research would have been accessible.
“Open API” is a term I didn’t encounter before and looked up. From what I understood, it is an initiative for standardizing how REST APIs are described by various vendors.

Evaluation:
As this paper is basically analysis of the existing text visualization approaches, though it is a good start in the area, it still needs a comprehensive study for the research to have any remarkable impact. The approaches and methods of expression that are included are by no means exhaustive and should be further explored to arrive at some significant results. The other aspect the paper needs improvement in is the analysis of the example visualizations that are mentioned. The paper just provides a cursory overview of each example that is included, it fails to discuss in detail the pros and cons of the particular approach to visualize the data.

Relevance:
The project our team is working on is in the area of literature, which basically deals with text data. I believe that studying various approaches and methods of expression in text visualization would be incredibly helpful for us to get an overview of which approaches we can use for ur project. Also, the example visualizations would be a place of inspiration for the visualizations our project to model on, especially the visualizations by Stefanie Posavec.