

Voltaire's Correspondences

Utilizing Visualization in the Mapping the Republic of Letters Project

Bugei Nyaosi¹

The project, Voltaire's Correspondence, aims to evaluate the utility of visualization in exploring distance relationships in the travel of letters within the Mapping the Republic of Letters project. The development of suitable visualizations will be discussed. Visualizations were created by using tools such as Tableau and Protovis to facilitate the analysis of data and to produce prototypes for final visualization. The process illustrates that visualizations are integral in the entire process of data analysis, and particularly in the Mapping the Republic of Letters project.

Background Information

Voltaire's Correspondences falls under the larger Mapping the Republic of Letters project, which studies the different correspondence networks that comprised the Republic of Letters. The Republic of Letters refers to intellectual communities in the late seventeenth and eighteenth century America and Europe that stretched across national boundaries. The Republic of Letters website notes that:

“The Republic of Letters was an intellectual network initially based on the writing and exchange of letters that emerged with and thrived on new technologies such as the printing press and organized itself around cultural institutions (e. g. museums, libraries, academies) and research projects that collected, sorted, and dispersed knowledge... Its scope encompassed all of Europe, but reached well beyond this region as western Europeans had more regular contact with and presence in Russia, Asia, Africa, and the Americas. In the sixteenth and seventeenth century merchants and missionaries helped to create global information networks and colonial outposts that transformed the geography of the Republic of Letters. By the eighteenth century we can speak of a trans-Atlantic republic of letters shaped by central figures such as Franklin and many others, north and south, who wrote and traveled across the Atlantic...”¹

The project aims to create visualizations by actually “mapping” these networks so as to study the spatial evolution of these networks over time. The Mapping the Republic of Letters database draws its information from the Electronic Enlightenment, an online collection of 57,685 letters and documents from a wide range of correspondences, and individually digitized correspondences.²

Voltaire

Without doubt, Voltaire is considered an emblematic figure in the Enlightenment and an integral part of the Republic of Letters. Voltaire was selected as a test case primarily because nearly 18,000 letters are attributed to him, which is a very rich and varied collection. Using Voltaire as a starting point enables us to compare Voltaire's correspondence with others, such as John Locke. This allows us to see whether there is an evolution in the practice of correspondence based on emerging patterns. Moreover, there is very little, if any research of this kind into Voltaire's correspondence.

Research Questions

One of the driving questions of the project is to determine the spatial extent of different correspondents' networks. Just how “national” versus how “global” were these networks? It was noted that looking at different statistical measures — such as average, median or standard deviation — of the distance traveled by letters has the potential to aid in understanding the nature of these networks. Examining the average distance letters traveled would also aid us in comparing the claims made about the Republic of Letters by its participants (e.g., that it stretched around the world, that it constituted an ideal republic) with the reality of a correspondence (e.g., Voltaire's correspondence seems strongly oriented toward a single country — France). Moreover, such analysis enables us to evaluate the function of correspondence — were correspondents such as Voltaire writing to acquire or disseminate information at a distance, or mostly writing to people nearby (whom they also, presumably, met in person)? This also led to another question that sought to understand how the destination of letters correlated with the location of senders. Did Voltaire simply write to people in the same location but due to his ever-changing location it appeared as if his network has a larger spatial extent than it actually had?

Ultimately, the project seeks to develop applications that can be standardized for use in the Mapping the Republic of Letters project and, in so doing, exemplify the use of visualization to interpret spatial information. Through this analysis, we can learn whether visualization, in this context, will show something new. While it is not within the scope of the project to analyze the content of individual letters in the database, there is a lot to be gained from analysis of the metadata i.e spatial information.

A well-established definition of information visualization is “the

¹Undergraduate Research Assistant, Stanford University Spatial History Lab



Figure 1. Mapping the Republic of Letters.³

use of computer-supported, interactive, visual representations of abstract data to amplify cognition”⁴ This visualization exemplifies how visualization can amplify cognition. It does an excellent job of clearly showing the letters in aggregation, and patterns are easily discernible. This particular visual compares the networks of Voltaire (highlighted in blue) and Locke (highlighted in orange). Immediately, one can point out the difference between the spatial extents of these networks. Voltaire’s is mostly centered in continental Europe while Locke’s is centered in the British Isles. While such patterns are almost immediately noticeable, it is nearly impossible to see the smaller distances that are obscured by the longer lines. Thus, it gives a false impression and a flawed

conclusion may be arrived at inadvertently. It would therefore be useful to also be able to explore all the distances fully, and to get a sense of the proportion of shorter to medium and longer distances. In pursuit of this improvement, we set out to explore other ways of visualizing the data we had in the database.

Methodology

The project database includes over 55,000 individual letters by more than 1,800 unique correspondents. Figure 2 is an excerpt from the database, showing some of the more important attributes

Document_send_date	Author_id	author_last_name	Recipient_person_id	Recipient_first_name	Recipient_last_name	Source_city	Source_lat	Source_long	Dest_city	Dest_lat	Dest_long
7-6-1751	20510	Voltaire	19296	Le Baillif	Ferriol, comte d'Argental	Potsdam	48.857	13.067	Berlin	52.516	13.377
8-7-1751	20510	Voltaire	18691	Charles Augustin	Ferriol, comte d'Argental	Potsdam	48.857	13.067	Paris	48.857	2.341
8-3-1751	20510	Voltaire	20070	Richard Gotthold Eapraim	Lessing	Berlin	52.516	13.377	London	51.508	-0.127
1-1-1752	20510	Voltaire	19368	Richard Gotthold Eapraim	Lessing	Berlin	52.516	13.377	Wittenberg	51.867	12.634

Figure 2. Excerpt from Project Database

of the existing data. It is important to note that we are working only with a sample of the database for which we have complete distance information.

The tools used in the process were Tableau and Protovis. Tableau is a data visualization software that enables anyone to easily create interactive data visualizations, dashboards and analytics. Protovis uses JavaScript and SVG to enable web-native visualizations. The aim of the initial visualization process was to provide error-checking using Tableau. It was necessary to geocode different locations so as to compute distance between points. This process tends to be prone to errors. Given the speed and ease with which visualizations can be built, Tableau is an excellent tool for quickly spotting anomalies or potential errors. In developing an application to visualize cities that Voltaire sent letters from, irregularities in the geocoding of cities were easily identifiable, as letters turned up in places from which Voltaire did not write. Figure 3 illustrates this feature.

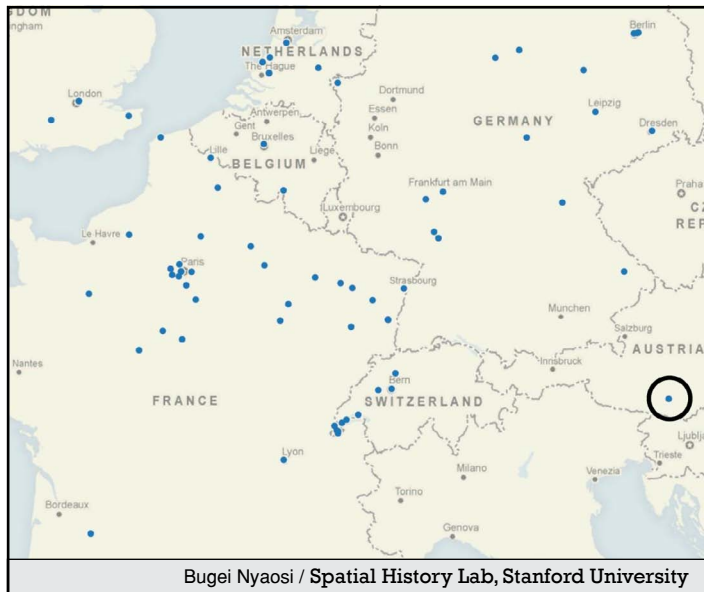


Figure 3. Map with anomaly. The highlighted point in Austria is suspicious, as we know that Voltaire did not stay in Austria. Therefore, it is highly unlikely that he sent any letter from that country. Tableau enables one to view the underlying data points, revealing that the particular city associated with the letter is actually in France. It was then relatively easy to change the point in the database to the correct geographic location. In this way, we were able to clean up the database and rectify any errors.

Analysis

One of the first charts that we produced plotted average distance traveled by letters against time. The thickness of the line varies with volume of letters. While this chart gave us a better understanding of the actual average distance letters traveled within a given year, the individuality of letters was lost. The average distance was also misleading as this could be the result of either a collection of medium-range letters or a combination of short- and long-range letters. The latter option is apparent in the spikes in average distance, as the relative thinness of the lines at those points suggests that a few high points are skewing the graph.

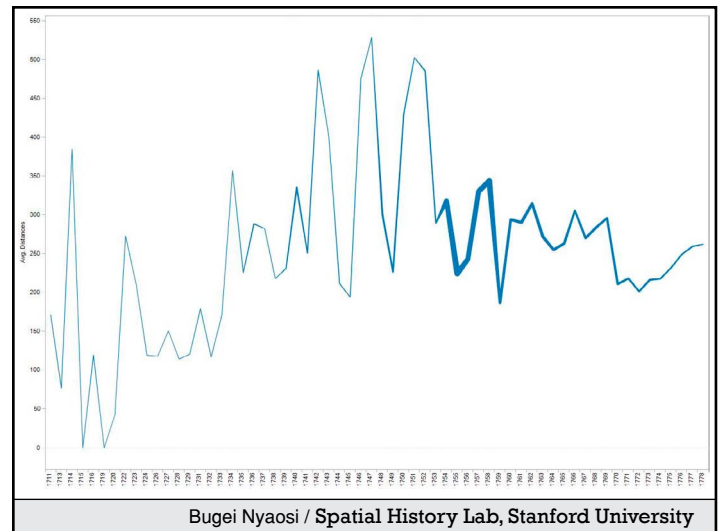


Figure 4. Line chart.

In addition, when we look at the standard deviation for the same points, it becomes quickly apparent that the averages are not the best measure to use, given the high values of standard deviation, confirming that indeed, these the distances are spread out over a large range of figures.

Year	Average Distance (Km)	Standard Deviation (km)
1728	114.3326457	198.0299513
1729	120.5642803	192.8598548
1731	179.2843423	106.3121786
1732	117.0231778	53.78884063
1733	170.9060604	102.376876
1734	357.1126316	311.6436336
1735	225.8177296	67.1278696
1736	288.6668411	398.2101029
1737	282.3689048	175.4007208
1738	217.9579592	108.4265777
1739	231.1623223	106.8381817
1740	335.5829242	344.130923
1741	250.7295525	95.39022019
1742	486.7496653	668.8064509
1743	401.0193887	201.3981125
1744	211.6047066	129.3416879
1745	194.2098363	330.5088048

Figure 5. Distances of letters sent by Voltaire.

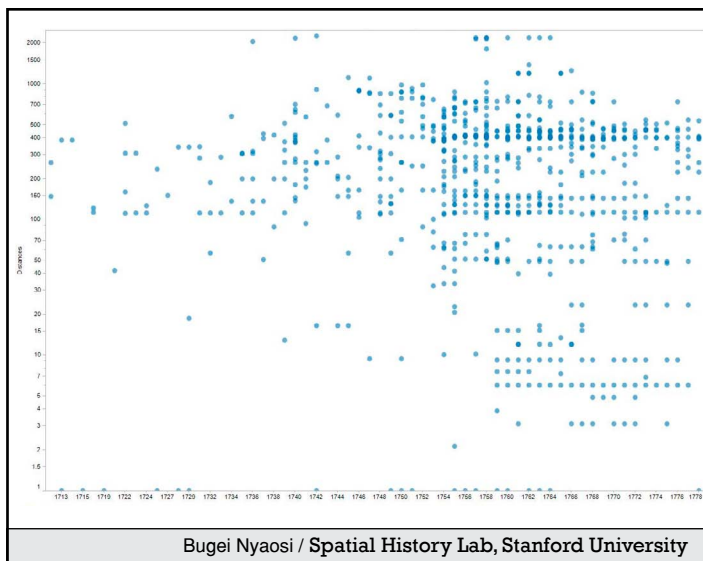


Figure 6. Scatter plot chart.

After further iterations, we constructed a scatter plot that would show each individual letter while still retaining the distance measure. This would give a better sense of the volume of letters and would also not just aggregate all distances — both long and short into a single point. Thus, it would deal with the problem of shorter distances being obstructed in the graph. A logarithmic scale was used to further enhance clarity. Hence, it would provide the ability to explore each letter individually. It is interesting to note that most of the distances were less than 500km, suggesting that most of Voltaire’s letters traveled within a relatively shorter radius than one would imagine. Moreover, there is a remarkable lack of shorter distances in the early part of his life. Why is this case? This could possibly relate to his relocation to Ferney in France, later on in his life where he stayed for a relatively long period of time as compared to his earlier years when he was constantly moving.⁵ While it may not be possible to answer this question based on this particular visualization, it points to the usefulness of this type of visualization for research purposes.

Despite the merits of the scatter plot, it did not provide a quick and easy way to get a sense of the overall distribution of distances. A histogram that broke down the data into “bins” was therefore incorporated. This particular chart is able to immediately give one a sense of the proportionality of various distances, constituting a good representation of the overall nature of the data. The data range show below, displays letters from Voltaire’s entire life span. It is very clear that a majority of letters are sent within 400km - 450km of Voltaire’s location, with almost all the rest below 400km. This confirms that most of Voltaire’s letters were sent over relatively short distances.

After experimenting with different ways of plotting distance traveled by letters, we settled on what we perceived to be a good visual representation of the data — a combination of the various different charts that we had initially come up with. This enabled us to capture various aspects of the data that would not be possible in a single chart. Thus, while we can see distances traveled by

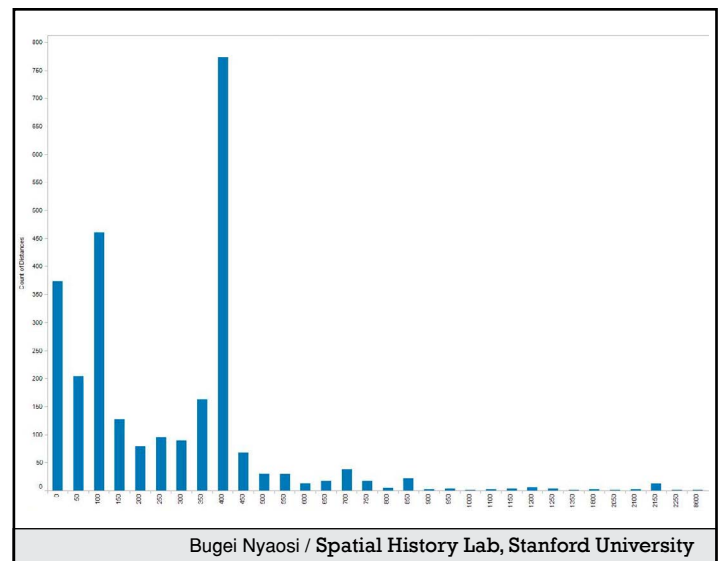


Figure 7. Bar chart.

individual letters, we can also see the average distance and distance distribution for a particular year or range of years. Interactivity was also incorporated into the visual, enabling filtering of data by year. This is especially useful in the case of the histogram, in that it allows one to view the data only for particular years.

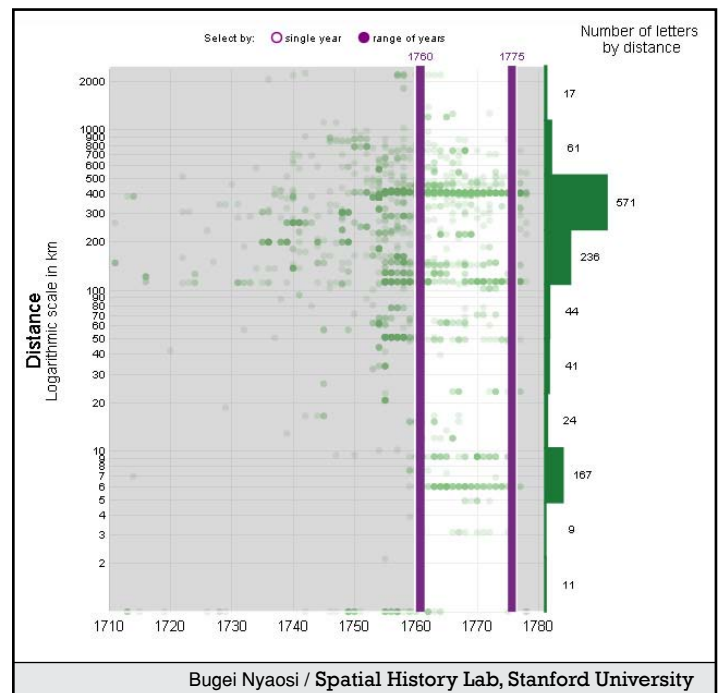


Figure 8. Final visualization. The final visualization, while functional, will only be truly useful when used within the Mapping the Republic of Letters project. It will then be possible to compare and contrast different correspondents. The scatterplot reflects the year sent and distance traveled by each of Voltaire’s 18,000 extant letters. Note that distance is plotted on a logarithmic scale, and that darker points represent more letters sent a given distance while more transparent points represent fewer letters. The histogram on the right hand side shows how the letters fall into buckets of distances. By dragging the purple vertical line(s) [\[in the online version\]](#), you can filter the histogram by a *single year* or a *range of years*.

Conclusion

The different visualizations created for Voltaire's Correspondence demonstrate the value of visualization in data analysis. Different charts enabled us to see different aspects of the same distance information. Some trends such as the relative short distance range of Voltaire's correspondence became more apparent depending on the chart. This shows that there is a lot to be learned from the different forms of visualizations in the Republic of Letters. The visualizations facilitate not only the answering of current research questions, but also the development of new questions.

End Notes

1. "Introduction," *Mapping the Republic of Letters*, 24 August 2010, <http://republicofletters.stanford.edu/>.
2. "Introduction," *Mapping the Republic of Letters*, 24 August 2010 <http://republicofletters.stanford.edu/>.
3. "Charts, Maps & Tools," *Mapping the Republic of Letters*, 24 August 2010, <http://republicofletters.stanford.edu/>.
4. Card S K, Mackinlay J D, Shneiderman B. *Readings in Information Visualization; Using Vision to think*. (Los Altos, California: Morgan Kaufman, 1999).
5. This observation is based on data from the database.

Author Information Correspondence and requests for materials should be addressed to Bugei Nyaosi bnyaosi@stanford.edu or Dan Edelstein danedels@stanford.edu.

Supplementary Information is linked to the online version of the paper at <http://spatialhistory.stanford.edu/publications>.

Rights and Permissions Copyright ©2010 Stanford University. All rights reserved. This work may be copied for non-profit educational uses if proper credit is given. Click here for additional permissions information.