Designing with hidden data

Stephanie Posavec

Many aspects of our world can be explained by mathematics and so, thus, can be broken down into series of numbers. The patterns found in numbers are often hidden to the viewer, as most examples of these patterns cannot be seen with the naked eye. These sets of numbers are found in places where we normally don't have the opportunity to notice the beauty of underlying rhythms, such as within a fleeting moment in time, or found lying dormant in an everyday object. By subjecting numeric 'codes' to a process that visualizes patterns and then converts them into form, we can help draw attention to these hidden elements of the world and let them be properly appreciated. By finding these undiscovered patterns and rhythms we not only create beautiful and complex visuals but also give the viewer a better understanding of the chosen subject matter. Also, it provides the opportunity to amaze and excite the viewer by representing the level of once-hidden complexity that pervades every aspect of our lives. For reasons using data to create visuals fascinates me, as I believe that the use of data-driven generative methods of arriving at design solution can be explored by anyone working creatively, regardless of their medium.

The album cover for the band OK Go showing visual derived from textual analysis by the lead singer.

For many years the use of data within design has mainly been found within the realm of information design. Information design has become more popular over the last few years as designers have become more interested in the visual representation of statistics—ranging from the mundane to the serious—that would normally

be communicated with words. The charts and diagrams found within this type of design are subjected to a strict set of rules: the information must be displayed in a clear and unbiased manner, where the aesthetic of the graphic does not obscure the data displayed within. Here, this design is created to be functional and useful: only the data is allowed to speak to the viewer, and anything subjective and non-quantifiable is irrelevant and should not be used. However, while there is a necessity for this type of design, there is also a place for visuals that use data in a warmer, more personal way. The visual outcome of the data can be used to communicate a subjective feeling or emotion without losing the ability to be analyzed and deciphered to understand the data within.

In order to use the data from the world around us to create a visual, it is first necessary to choose which data to gather from the chosen subject. Anything that is quantifiable and can be converted into a numeric "code" can be used to create a visual. After selecting and gathering the information to be visualized, the next step is to create a system of specific aesthetic choices that will occur in varying intensities according to the numbers found in the data. The data is processed using this system, and this process is what creates the final visual outcome. Here, what you see is what you get: any quirks and idiosyncrasies in the final form should be appreciated as part of the unique quality of the selected data-set. To remove data in a quest for visual perfection would ruin the integrity and meaning found within the visual.

I use this method when producing my work, where I enjoy exploring the underlying patterns hidden both in text and within a writer's personal writing style. I first became interested in the visual analysis of text when working on my final project on the MA Communication Design course at Central Saint Martins College of Art & Design, London, in 2006. My project was entitled "Writing Without Words": that is, I wanted to explore new visual methods of representing text unlike its traditional form of lined pages within a book. I also wanted to find a way to visualize the differences in writing styles between different authors, though I chose to focus the bulk of my analysis on the novel On the Road, by Jack Kerouac. For this project I chose three different types of data to measure and represent in a different form: a sentence's rhythm, sentence length, and a text's structure. The outcome of this project was a series of posters and book, and completing this first project helped me evolve the process for how I produce text visualizations today.

While reading the text that I have chosen to visualize, I become excited by visual themes, words, or the author's writing style, and this excitement helps me select the most interesting data from the text that would potentially resonate with the viewer and help them see the text afresh. After I make these decisions, if I have time, I analyze the text by hand using various colored markers and pens. I painstakingly analyze the text according to criteria such as syllables, parts of speech, words

Posavec's hand-analysis of the text The Influence of the Blue Ray of the Sunlight and of the Blue Colour of the Sky *by General A. J. Pleasanton, part of the analysis of the parts of speech in an excerpt from the OK Go album's namesake.*

per sentence, or themes found within the text. By the end of this analysis, the text is reduced into numbers and single letters: the text's unique code. Of course, using a computer program to analyze the text would be more efficient but the handmade, craft feel of this process appeals to me. Actively engaging in the data-gathering process is an important part of the project:

for me, the effort involved adds more meaning to the final piece. When finished with analyzing the text in this manner I end up with a stack of hand-notated pages of color, symbols, and forms: a type of handmade data visualizations.

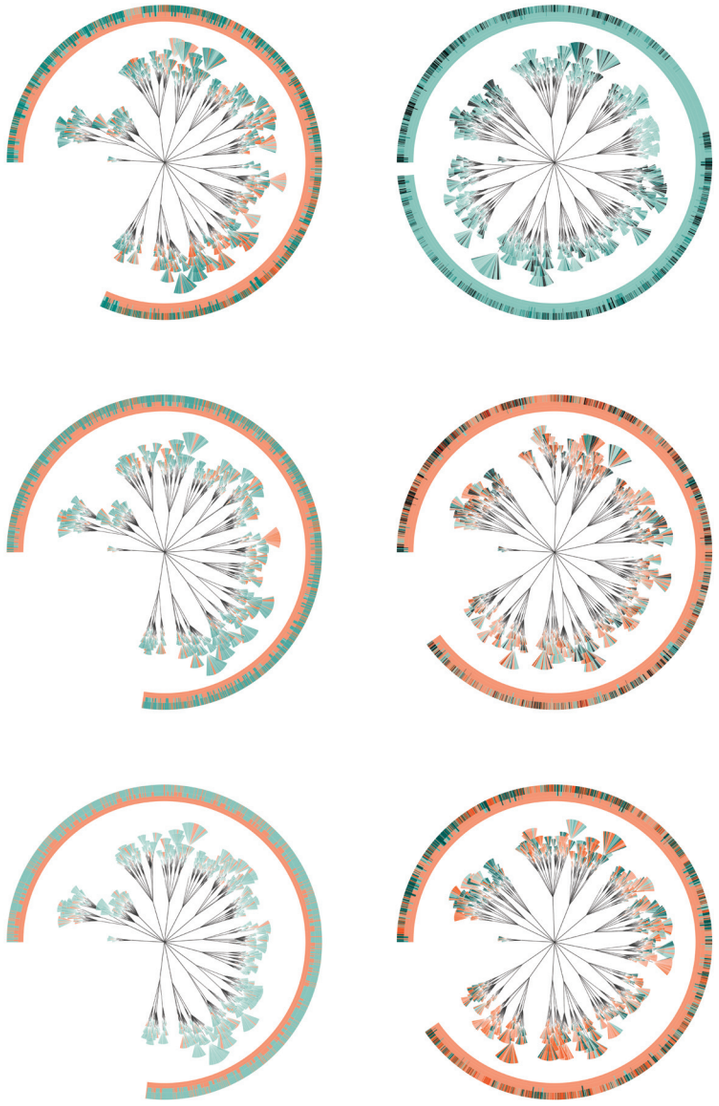
Using this data, I focus on the visuals: how can I use this newfound pattern to create a form that looks beautiful but still can easily communicate its information to the viewer? I will often search the text for visual metaphors that will provide the basis for the visualization's aesthetic, or will try to shape the data into a form that represents my personal reactions or perceptions of the text. For my "Writing Without Words" project, I wanted to visualize the structure of a novel. I also wanted to communicate how a long-loved piece of literature feels warm and vibrant to me in a way that is not visually communicated when the text is in book form. I enjoy how both books and living beings are cellular in that both are composed of tiny components that create a larger entity. I decided to use a simple tree diagram to visualize the structure of Part One of On the Road: here, Part One divides into chapters, chapters divide into paragraphs, paragraphs divide into sentences, and sentences into words. By using a tree diagram to represent the structure of the text, I was able to create an organic, plant-like structure that both displayed the data for closer analysis and also provided a visual metaphor for my perceptions and emotions about literature. Here, the merging of emotion and the visual display of information can be combined to create a stronger, more meaningful visual.

The reason that I am drawn to producing visuals in this manner is that I believe it adds meaning to the work for the viewer. The viewer's knowledge of how the form was created suffuses a visual with a hidden sense of meaning and importance. Much like writing messages on a wall before it is painted or burying a time capsule in the ground before constructing a building, this very knowledge has a talismanic quality, saturating the form or space with meaning, and investing it with intention and importance. However, while this method can be used to produce a visual, and while the viewer or user can be made aware of the data-set that was used to generate the visual, it is vital that there is an explanation that is made available so it is possible to make new insights into this data while also appreciating its form. For me, there is no point in using this method unless there is an access point where the viewer or user can explore the data further.

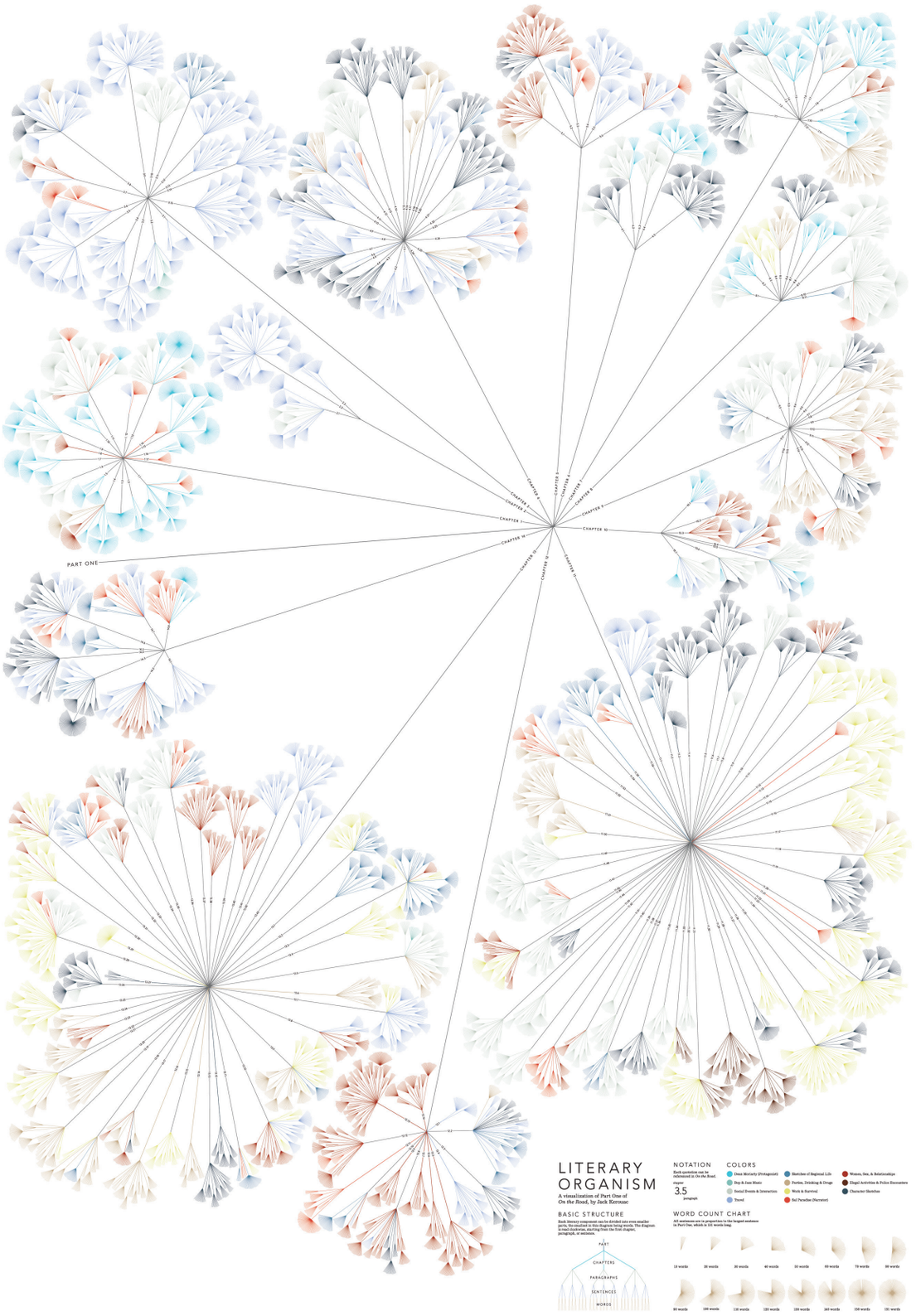
This method of working can be quite personal and the goal of this process is equally to enhance my understanding of a particular subject as much as it is used to communicate new ideas to the viewer. When working alone, I tend to move inward with ideas, which hinders the possibility of connecting with an outside world. I find that the act of collaboration has been a useful tool to ensure that what I create is relevant and useful to a wider audience. Working with a wide variety of people from different backgrounds ensures that I begin to think about the project's potential function and purpose instead of focusing on how the process benefits me alone.

An example of how this process has been used in a collaborative manner is in the project "Entangled Word Bank," a collaborative project between myself and Greg McInerny, an ecologist working at Microsoft Research in Cambridge, England. Greg initiated this project as part of his research and invited me to work with him to explore ways of visualizing the text of On the Origin of Species by Charles Darwin. We created images that both represented the insertions and deletions of text over the six editions and also visually alluded to the study of plant specimens and methods of botanical dissections. Because of the quantity of text within this project, I had to accept that I wouldn't be able to analyze the text and produce the final artwork by hand, but Greg would instead create programs to both analyze the data and draw the final visuals. Effectively, my personal analysis and connection to the text was broken: I had no hand in creating the actual analysis and visuals; rather, the computer program completed this for us. However, the project became stronger for it: by creating a program that could theoretically be used to analyze any text in the English language, we created a set of visuals that could provide a function if refined further. Here, collaboration elevated this process to have the potential to be a useful tool for text analysis instead of only being aesthetically-pleasing.

I have also found that these projects require the client to be involved more deeply than normal in the production of the project, bringing forward a greater sense of collaboration in the client/designer relationship. For this type of collaboration, the client's analysis of a topic is often needed to produce the visual, and will require



All 6 edition diagrams from Origin of Species: within the diagram, chapters are divided into sub chapters as in Darwin's original text, and these sub chapters are divided into paragraph 'leaves'. The small wedge-shaped 'leaflets' represent sentences. Each sentence is colored according to whether the sentence will survive to the next edition (blue) or whether it will be deleted and not be within the next edition (orange).



'Writing Without Words.' A literary organism poster visualizing the data contained within the text for part one of Jack Kerouac's On the Road.

a level of effort that goes above and beyond the average client's input. Greg McInerny and I also worked together to create the artwork for Of the Blue Color of the Sky, the third album by the band OK Go. The title of the album was taken from a scientific text written in the nineteenth century, and the band wanted us to create visualizations of both this text and their song lyrics for the album artwork. For this project, intensive collaboration between designer and client was vital to ensure that we were capturing the themes and information that the band found most interesting. Lead singer Damian Kulash Jr. carefully analyzed the book and song lyrics sentence by sentence according to twenty-five themes of his choosing. As this was his personal, subjective opinion about the text, Greg and I wouldn't be able to closely replicate this data if we analyzed the text ourselves. This data was used to create the artwork for the album cover. Perhaps the knowledge that the lead singer's analysis was the basis for the album artwork will increase OK Go's fan base's interest and excitement in the project.

Whether a designer chooses to work alone or with other people, this method of creating visuals is something to be explored when experimenting visually. Generative visuals created from datasets can be applied to generate patterns for surface design or inspire visual form in three dimensions, such as sculpture, interiors, or buildings. This way of working should be considered amongst other methods in a designer's "toolbox" as a different way to arrive at a design solution for a particular project. Much like photography is often used to highlight the beauty of the mundane or the unnoticed, data-driven art highlights the intangible beauty of the numbers inherent within the tangible elements of this world: an area of rich visual possibility that can be used to arrive at new ideas.