

Participatory Visualization with Wordle

Fernanda B. Viégas, Martin Wattenberg, and Jonathan Feinberg



Abstract—We discuss the design and usage of “Wordle,” a web-based tool for visualizing text. Wordle creates tag-cloud-like displays that give careful attention to typography, color, and composition. We describe the algorithms used to balance various aesthetic criteria and create the distinctive Wordle layouts. We then present the results of a study of Wordle usage, based both on spontaneous behaviour observed in the wild, and on a large-scale survey of Wordle users. The results suggest that Wordles have become a kind of medium of expression, and that a “participatory culture” has arisen around them.

Index Terms—Visualization, text, tag cloud, participatory culture, memory, educational visualization, social data analysis.

1 INTRODUCTION

It is the dream of anyone who aspires to create visualizations for a mass audience: watch as your work is viewed millions of times on the Internet, see it on the nation's biggest newspapers, have it used by scientists, kids, and politicians. Jonathan Feinberg had precisely this experience when he created the Wordle web site (<http://www.wordle.net/>), a tool for making "word clouds". In the nine months since its launch in June 2008, over 600,000 such word clouds (or "wordles") have been created—a pace of one new wordle every 10 seconds. How and why did the Wordle site capture the imagination of so many people? And what can it tell us about the place that data visualization might take in mass culture?

Wordles are close relatives of tag clouds, encoding word frequency information via font size. Despite the shared DNA, wordles look remarkably different from regular tag clouds. They are striking graphic statements. Beyond presenting word frequency data, wordles revel in the possibilities of color, typography, and composition.

In turn, users seem to revel in the possible applications: scientists wordle genetic functions, fans wordle music videos, teachers wordle literary texts, spouses wordle love letters, kids wordle their thoughts, and journalists wordle political speeches. Wordles have made their way into corporate PowerPoint slides and houses of worship; they have adorned T-shirts, magazine covers and comfy cushions. Wordles have graced YouTube videos, homemade gifts, posters, scrapbooks, birthday cards, and valentines.

To be sure, some people create wordles to engage in data analysis, but they are a minority. Indeed, there is evidence that a significant number of users do not even understand what the graphics are encoding.

(Whether wordles qualify as an effective analytic tool is up for debate and we discuss this question below.) One might conclude on these grounds that wordles are “mere” entertainment, and that their popularity is no more interesting than that of a hit movie or a cute Facebook application. Indeed, the Wordle site describes itself as “a toy.” Yet such a simple explanation does not capture the complexity of Wordle’s usage. While people do report making wordles “for fun,” they show an impressive level of creativity in using its output.

This paper examines the source of Wordle’s popularity: what makes it tick and what draws people to it. By studying spontaneous usage of Wordles on the web, and by running a survey with thousands of respondents, we start to unravel the Wordle riddle. We suggest that a key message of the Wordle phenomenon is that scientific-style data analysis is not the only *raison d’être* of visualization tools. Our results suggest that Wordle usage may be viewed as a component of “participatory culture” (in the sense of Jenkins [5]): a cultural system in which viewers are also producers and remixers, and where visualization serves as much as an authoring tool as a method of analysis.

1.1 Related work

Our work relates to several streams of investigation in visualization. Tag clouds, simple visualizations that display word frequency information via font size, have been in use on the web since 1997 [13]. In recent years, several sites have arisen that create tag clouds from user input. These range from single-purpose tools such as TagCrowd [7] or The Tag Cloud Generator [10] to more general visualization sharing sites such as Many Eyes [12]. Researchers also have looked at ways to improve on traditional tag cloud layouts. Seifert *et al.* [7], for instance, created an algorithm to efficiently place rectangles into arbitrary polygons, and used that to create more tightly packed word clouds, while Gambette and Veronis recently introduced the “Tree Cloud” [TK]. Wordle fits directly into both these traditions.

One of the themes of this paper is the importance of non-analytic uses of visualization, an observation that others have made as well. The

- *Fernanda Viégas, Martin Wattenberg, and Jonathan Feinberg are with IBM Research, E-Mail: {viegasf, mwatten, jdf}@us.ibm.com.*

Manuscript received 31 March 2009; accepted 27 July 2009; posted online 11 October 2009; mailed on 5 October 2009.

For information on obtaining reprints of this article, please send email to: tvcg@computer.org.

communicative and illustrative aspects of visualization are well-known. More recently, the genre of “social visualization” [3] seeks to convey social information in a non-scientific form, much as the ambience of a café reveals information to a viewer. In [4] Hearst specifically points to tag clouds as a form of social visualization.

The notion of “casual infovis” [6] is also directly relevant to our investigation. Wordle qualifies as a casual infovis system, since it may be used by non-expert users to depict personally meaningful information. However, we will argue that the usage scenarios that we see with Wordle go beyond the definition of casual visualization in several ways, especially in the many cases where it seems to function as a remixing or authoring tool.

The level of discussion that has occurred around Wordles, and its public presence on the internet, suggest relations with work on the Many Eyes system [12] and the Name Voyager [14]. These two systems were heavily oriented to interaction and data analysis. (In addition, the NameVoyager was built for only one data set.) In contrast, Wordle is tuned for rapid authoring and display. As we’ll discuss, this difference in emphasis led to significant differences in usage centering on the theme of user creativity.

In fact, as we’ll argue, the response to Wordle hinges so strongly on the notion of user creativity that it may be fruitfully viewed in the framework of Jenkins [5], who discusses “participatory culture.” In his definition, this is a culture that includes “low barriers to artistic expression and civic engagement, strong support for creating and sharing one’s creations, and some type of informal mentorship.” It is a culture of convergence where consumers are encouraged to seek out new information and make connections among dispersed media content. This concept contrasts with older notions of media spectatorship where producers and consumers occupied separate places. Classic examples of participatory culture include “fan fiction,” video remixing, and, more recently, knowledge communities that spring up around shows such as “Survivor” and “Lost.” These are all instances where viewers take existing content (images, text, video, a story plot), mix and match pieces to create new artifacts. Such artifacts may take the form of video clips, a new novel, a sequel story plot, or a collage of images. A lot of times these activities are carried out as playful pursuits, other times they take on a more serious undertone—for example when people use remixing to protest government actions, to support political candidates, or to call attention to a worthy cause. Participatory culture encompasses the work and play that users perform in new media systems and we propose that Wordle users are a part of it.

2 ABOUT WORDLE

Before delving further into the question of Wordle’s usage, we describe the visualization technique in more detail, as well as the user experience of wordle creation. The appendix details the Wordle layout algorithm.

2.1 The Visualization

The motivation behind the Wordle site was to provide a more pleasing version of a traditional tag cloud. Tag clouds have typically used rows of alphabetized text, with the juxtaposition of different font sizes leaving gaps of white space in an awkward “ransom note” effect; typefaces are usually limited to standard browser fonts.

Wordle seeks to fix these typographic problems. The Wordle layout packs words tightly, requiring only that the actual glyphs of the word do not intersect—so that, for example, a tiny word can appear in the middle of the space of a letter O in a larger word. This contrasts with earlier methods, such [7], whose placement algorithm approximates words by their bounding boxes

Wordles also provide users with a broad set of visual parameters to play with: a large set of typefaces and several pre-fabricated color schemes. The range of typographic possibilities is in line with a word processor, a rarity for a visualization. Wordle also breaks away from the pure horizontal alignment of conventional tag clouds, giving the option to place words vertical or even diagonally.

2.2 The user experience

The process for creating Wordles is straightforward: users go to a web page, paste a piece of text into a text area, hit a button, and get a word cloud.¹ The initial cloud is generated with randomized visual parameters. From this starting point, users may customize the look of their creation. This process is critical to the user experience, so we describe it in detail. The entire process is designed to be simple and playful, imparting a feeling of control over the creation process.

Font

As with a standard tag cloud, wordles use font size to represent word frequency. Unlike many conventional tag clouds [12], the font size is linearly related to frequency, rather than the square root. This choice was made by wordle’s creator because the resulting compositions seemed more dynamic.

Wordle offers 34 fonts, ranging from classic faces to more entertaining and whimsical choices. Ubiquitous fonts such as Lucida Sans live side by side with less well-known, quirky options like *Boopee*, *Loved by the King*, and *Alphabet Fridge Magnets All Caps*.

Color

Wordle provides 12 palettes, from black on white to fresh-sounding combinations like *Asparagus*, *BlueSugar*, and *ChilledSummer*. Users may also create their own color combinations. To add subtler nuances, users may add different degrees of random variation in the brightness and hues of the base palette colors.

Content/Language

Common words like “the” or “of” usually hold little interest for users. Wordle removes these “stopwords,” which otherwise clutter the display. The software has lists of stopwords in 26 languages, and by default uses whichever list has the most words in the input text. Users may override the default, choose to leave common terms in the visualization, or remove additional words. Further options let users include or remove numbers, and force upper or lower case.

Layout

A Wordle layout is defined by several parameters: the number of words shown, the angles to draw text, whether alphabetical order should affect placement, and whether the overall shape should be “blobby” or “straightish.” Unlike color, typeface, or language, users have only indirect influence on the layout due to randomness inherent in the packing algorithm. Nonetheless, users can influence the composition, making wordles that range from intricate clouds with hundreds of words to minimalist pieces with just a few terms.

Sharing

The Wordle site supports sharing in a number of ways. Users may save a new wordle to a gallery, giving it a unique URL and allowing others to see it in a “browse” page. A “print” option generates a high resolution version of the visualization. To provide legal protection for sharing, every Wordle is licensed under a Creative Commons Attribution license which allows free usage for commercial or non-commercial purposes.

3 STUDYING USAGE: METHODOLOGY

The Wordle site has seen heavy usage.² Logs indicate an average of 14,000 visitors per day. Unlike many short-lived internet phenomena, traffic has trended slowly upward over time, with an average of 11,000 daily visitors in August 2008 and 22,000 in February 2009. One way to get a sense of the traffic is to compare with web sites of city newspapers.

¹ Wordle also accepts two additional kinds of input: (1) the URL of any blog, blog feed with Atom/RSS feed, and (2) a del.icio.us user names.

² Although the Wordle site launched in June 2008, detailed traffic logs are available only from July 22 onward. The statistics in this paper apply to the period from July 22, 2008 through March 4, 2009.

Portrait of the candidate as a pile of words



Fig 1: Boston Globe's Wordle of the Republican and Democratic campaign blogs.

According to compete.com, Wordle receives a level of traffic similar to the daily papers of Reno, Nevada, and of Tallahassee, Florida. The question, therefore, is what are all these people doing? Why do so many people find value in Wordle, and what sort of value is it?

Visitors to the Wordle site leave few traces other than the Wordles they make. They don't have to register or log in, and the site does not even keep a copy of the original text used to create the wordles. Such features are a boon for users, who can quickly create word clouds, but it means researchers have no information about a Wordle's creator or about the text it was based on.

To understand why people create Wordles and how they use them, we therefore took on a two-pronged study: "Wordles in the wild" and "Wordle survey." The former is a look at spontaneous usage via web-based accounts of Wordle, and the latter reached out to users for a more structured view of people's experience in creating Wordles.

3.1 Wordles in the Wild

The Web is Wordle's birthplace and a natural locus for learning about its use. We found examples in several ways. In some cases we simply stumbled across Wordles on high-profile sites. We also did a Google search for "wordle" during the first week of 2009 and looked at the first 500 links (Google claimed to have nearly one million results total). These links led to web sites that mentioned Wordle, often with detailed descriptions about user motivations and experiences. Based on this reading, we performed targeted searches to learn more about particular use cases. For instance, seeing evidence that many teachers used Wordle, we performed a search for "wordle" and "education" to learn more.

3.2 Wordle survey

To complement the search-based investigation, we asked Wordle users directly about their experience. We created an online survey and added a link to this survey from the home page of the Wordle site. The survey consisted of 25 questions: 18 multiple-choice and 7 freeform text. The questions covered users' motivations for creating Wordles, plans to share their creations with others, editing behavior, reactions to their Wordles, and an explicit comparison with tag clouds.

4 RESULTS

4.1 Wordles in the Wild

4.1.1 Themes

Because Google search hits are ranked by a rough measure of authority, the "Wordles in the Wild" approach provides a glimpse at the most prominent contexts for Wordles on the Web. Our goal was not to determine a quantitative breakdown of the proportions of different uses—that would probably be impossible, given the difficulties of obtaining a truly random sample—but rather to identify and learn more about some of the popular uses for Wordle.

The search results reveal a diversity of both usage and users. They included sites ranging from newspapers and magazines to educational pages, from poets' personal blogs to music and community sites. Children and adults alike describe using Wordles to goof off or to perform tasks for work or school, and to view texts ranging from Twitter feeds to books. From this eclectic set of scenarios, three themes appeared repeatedly.

4.1.2 Usage in the mainstream media

During the American presidential campaign of 2008, Wordles could be seen in many media outlets. Major newspapers and magazines added political Wordles to the national debate around the election. A favourite way to use Wordle was to **compare the rhetoric of the two major parties**. For example, both the Washington Post and the *Boston Globe* featured a side-by-side comparison of the Democratic and Republican candidates' blogs (fig 2), with the takeaway being that "Obama" was, by far the most prominent word on McCain's campaign blog.

WIRED Magazine took comparisons further; their online blog created a gallery of eleven Wordles depicting speeches from both major party conventions. After the election, political usage of Wordles continued. For example, the change.gov site created by the incoming Obama administration used a Wordle to portray citizen feedback on health care reform.

Media usage of Wordles was not limited to politics, however. **Often Wordles were used to distill user-generated content from the web**. The *Guardian* wordled the outlook for 2009 as defined by Twitter users, for instance, while a *WIRED* magazine piece entitled "Mourning the Internet Famous: Randy Pausch's Distributed Funeral" used a Wordle to illustrate the top 20 most used words in comments from Tributes.com on the professor's death. In some cases the media simply drew attention to the tool, as when television personality Rosie O'Donnell spotlighted the Wordle site in her blog.

As with general traffic to the site, these examples didn't represent a passing spike of attention. Usage by top international media outlets continued long after Wordle was launched. In February 2009, wordles appeared on the web sites of the *London Times*, *USA Today*, the *Australian Globe & Mail*, and *US News and World Report*.

4.1.3 Personal Usage

Many users created wordles of personally meaningful data. Sources range from online media (blogs, Twitter feeds, etc.) to poetry and song lyrics. People often write about these creations on the web: A Google search for "+wordle" +lyrics³ yielded more than 42,000 results as of March 2009.

One common pattern is for a group of people to find each other online and create a set of related wordles. For example, one blogger created a Wordle of last names contained in a personal genealogical database. He posted his wordle and gave step-by-step instructions for how to export the data from a popular genealogical software package. In response, other bloggers picked up the idea and continued the process. On a different blog, for instance, titled "*Saturday Night Genealogy Fun - do a Wordle*," a fellow enthusiast posted his own wordle and asked readers, "*Can you make something else really creative or pretty?*" In the comments section half a dozen people shared their creations.

³ In our search queries we put *wordle* in quotes, to prevent Google from automatically "correcting" it to *world* or *word*.

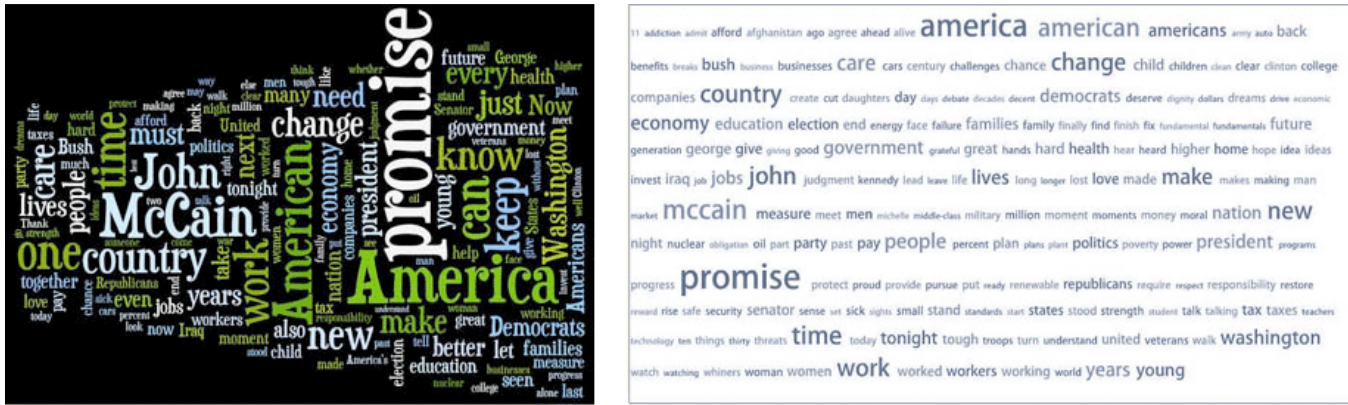


Fig 2: Wordle vs. Tag Cloud of Barack Obama's speech at the Democratic Convention in 2008.

In other cases, people gathered around the idea of making wordles. On the Flickr photo sharing site, a search for “wordle” yields more than 4,000 results as of March 2009. Some of the images are part of special “photo pools.” These range from the 135-person general-purpose “Wordle” group to the 5-person “namethatwordle” group (where people wordle lyrics and viewers to guess the song.)

Part of the excitement around Wordle seems to lie in the ability to take visualizations beyond the Web. Because the site lets users export a high-resolution version of their creations, savvy users can put wordles onto T-shirts, cards, and other physical objects. One blogger provided instructions on how to go from the wordle web site to a finished T-shirt. Another blogger wrote, “*What a fun Valentines Gift, click here. When you get there click on create. I am going to do it for my girls tomorrow. Print it out. Embellish it a little. Frame it. Might even do it for my hubby.*” The use of Wordles for romantic mementos was common. In one online video, the narrator described how he “got boyfriend points” when he made a Wordle out of love letters written to his girlfriend. Another woman, on the “Wedding Bee” blog, described making a Wordle of her wedding vows.

4.1.4 Usage in education

Educational uses of Wordle constituted a third theme. A striking number of sites gave extensive tutorials on how to use the tool in the classroom. (Perhaps one might expect this from an audience of teachers!) A Google search for “wordle” and “education” brings up many results, including videos, lists of exercises, and essays. A blog post titled “The Top 20 Uses for Wordle,” suggested using Wordles for classroom introductions and to display Shakespearean sonnets.

An example lesson plan comes from a teacher who mailed Wordle's creator and allowed him to reprint her description on the Wordle blog. She taught spelling by creating Wordles of new vocabulary words, and then quizzing students on various aspects of the displays. After multiple iterations, the test results “blew her away”. She was also impressed that the exercise engaged even less motivated pupils—as she put it, “*the last of the hardcore "I ain't doing it" students had a go and surprised us both.*”

4.2 Wordle survey

To run our survey, we deployed a set of pilot questions on January 20, 2009. After 2 days we collected 720 responses. Based on this round of responses and feedback from participants, we revised our questions and deployed a final survey on February 5, 2009. The survey ran for one week and generated 4,306 responses, which we analyze below. It is important to note that our survey respondents were self-selected, so that our findings must be interpreted with caution—they are at best a rough guide to what an “average” user experiences.

Two main themes emerged from our results: the importance of design and that the Wordle site works like a creation tool. In addition, a significant portion of participants turned out not to understand what font

size meant in Wordles. We discuss these findings in detail in the following sections.

4.2.1 Respondent Demographics

Female respondents outnumbered males by a two-to-one margin (67% female versus 33% male⁴). Wordle attracts people of all ages: 22% of respondents were under 20, 20% between 21 and 30, 19% between 31 and 40, and 39% above 41. Educators and students make up the largest group of users in Wordle: 29% of respondents were students (12% were high-school students) and 17% were teachers and professors. No other type of occupation constituted more than 6% of the total.

4.2.2 Design/Typography

Design and visual appeal were overwhelmingly cited as a reason for users' interest in Wordle. Not only were wordles attractive, the design inspired users to engage with the visualization in creative ways. When asked why they create Wordles, participants made it clear that “beauty” directly affected how they related to and ultimately used Wordles:

I think the images are beautiful and I'm looking for a way to get them off the screen and onto fabric so I can make some cushion covers

I wanted to have an artful, beautiful way to remember good quotes, good ideas.

They were entertaining and beautiful, a great way to express myself. They were helpful, pleasing to the eye, a pleasure to make and a pleasure to view. I have used Wordles in cards, for school, and mostly for my own enjoyment.

Because Wordles are closely related to tag clouds, we wished to compare people's reactions to both techniques. Do users prefer the clarity of tag clouds or the more free-flowing layout of Wordles? Does rotated text bother users? To find out, we asked questions about four Wordle-tag cloud pairs—each depicting the same text: speeches delivered by candidates at the Democratic and Republican Conventions of 2008 (Fig 2). When asked which representation was more effective, 70% of participants felt Wordles were more effective compared to 11% who found tag clouds outperformed Wordles (19% of respondents felt both representations did equally well). Participants' preference for Wordle can be broadly grouped into three main categories: **emotional impact, attention-keeping visuals, non-linearity.**

⁴ Due to the large number of responses, the statistical margin of error is less than 1%. Since the sampling is not random, however, these numbers should be taken only as a general guide.

Emotional impact. Emotional resonance seems to be a key reason for preference of wordles over tag clouds. Explaining why they thought Wordles were more effective than tag clouds, users gave reasons such as:

[Wordles have] *More emotional emphasis, colours and layout enhance the meaning.*

[Wordles are] *Colorful, more visually interesting, more of an emotional response and connection with the viewer than the tag clouds.*

[Wordles are] *More engaging. The colours help your eyes to move around and take in different words. The design of Wordles gives them more visual and emotional impact.*

Attention-keeping visuals. Not only are Wordles good at catching people's attention, but participants explained that they might examine Wordles longer than they would tag clouds:

Wordles are more effective because they are more attractive and interesting; I would be more inclined to read them.

They are more visually engaging and stimulate further inspection. They are more interesting to look at, and therefore you will look at them longer/more closely.

They look like art and make you look twice, then all of a sudden you are reading, thinking, feeling good, learning

Organic/Non-linear. Wordles subvert some basic rules of text legibility as they break line continuity, show text sideways, on a diagonal, etc. Users seem to revel in the non-linear side of Wordles:

The wordles are like the way we think--not always linear. sometimes one word runs right into another. more connections can be formed.

The words on the Wordle stand out more to me - I am a random, creative thinker, not a linear thinker. It feels more organic, and human.

There is a clarity in the tag clouds -- if you want to know whether a word was used (and how often), it's easy to locate. But there is a coherence, in a sense, to the Wordles; they feel organic, the way real speech feels.

Drama, central words stand out, much more interesting to the eye than the linear presentation, color enhances the presentation - it becomes a statement and a piece of art - linear is flat and toneless.

87% of respondents used Wordle's customization capabilities—either by trying different layouts or specifying combinations of font and color. Asked why they had edited their visualizations, users generally mentioned a desire to make the visuals match the content of the text. Some respondents also wanted the visualization to match either a specific identity (e.g., a corporate brand) or someone's personality (their own or the person for whom they had made the Wordle). The frequency of edits speaks to users' interest in experimenting with typographical arrangements that fit their needs.

4.2.3 Creation Tool

The fact that users make ample use of font, color, and layout choices points to a second theme: a feeling of creativity in using Wordle. At least within our survey sample, wordles were not passively consumed: 76% of our respondents said they had personally produced one. But what sorts of text were people visualizing? Most respondents used texts they were already familiar with, with the majority wordling their own writing (table 1).

Table 1: Familiarity with text in Wordle

| How familiar were you with the text before making the Wordle? | Percentage |
|---|------------|
| I wrote it | 57 |
| I read it many times | 19 |
| I read it once | 9 |
| I skimmed it | 6 |
| I had never looked at it before | 7 |

Asked about their experience in using wordles, 88% of people said they felt creative (table 2)⁵. This result held up across the board, even after controlling for age and gender. Most people (66%) also said they felt an emotional reaction to these representations.

Table 2: Users' experience with Wordle

| | Agree % | Neutral % | Disagree % |
|---|---------|-----------|------------|
| I felt creative | 88 | 9 | 4 |
| I felt an emotional reaction | 66 | 22 | 12 |
| I learned something new about the text | 63 | 24 | 13 |
| It confirmed my understanding of the text | 57 | 33 | 10 |
| It jogged my memory | 50 | 35 | 15 |
| The Wordle confused me | 5 | 9 | 86 |

Although 46% of respondents agreed that they had used their Wordle as an analytic tool, even more people cited decoration, illustration of a point, or simply using it for fun.

Table 3: How respondents used Wordles

| | Agree % | Neutral % | Disagree % |
|---|---------|-----------|------------|
| I was just trying it out for fun | 81 | 9 | 9 |
| To decorate websites, presentation, print, etc. | 51 | 18 | 29 |
| To illustrate a point I was making | 50 | 22 | 27 |
| I used it as a memento or souvenir | 47 | 19 | 32 |
| I used it as an analytic tool | 46 | 22 | 31 |
| I used it as a teaching tool | 37 | 23 | 39 |
| I used it as a gift | 30 | 23 | 46 |
| I didn't use it for anything | 18 | 17 | 64 |

Finally, the survey included a freeform question where respondents could elaborate on why they had created word clouds. Four main themes emerged: fun, creativity, educational purposes, and gift giving. Data analysis also figured in the answers, though not as prominently. The quotes below give a sense of people's motivations for creating Wordles.

For Fun

For fun, and to use in art

It was fun and we did it for math.

I created the my first Wordle on an assignment for school. I loved it, so I started using it for pleasure. The images created are very cool and artistic and fun ways to express yourself.

Creative

To use as a background in a fabric and paper book cover.

I am using it as backdrop to teachings at church.

I am a digital scrapbooker and I LOVE the way that Wordle creates titles (i.e. descriptive titles) for my layouts.

Educational

I am a teacher. Wordles help me introduce topics to pupils by allowing them to see what a new topic is about. In academic writing, I use Wordles to check for excessive use of any word or phrase.

⁵ Our survey provided users with a five-point Likert scale that ranged from "strongly agree" to "strongly disagree." For tables 1 and 2, we report agreements as the aggregation of "strongly agree" and "agree" answers. "Disagree" and "strongly disagree" are reported under "disagree."

Teaching my students about synonyms. Each student was given an 'overused word' and a thesaurus. Their task was to find at least 4 synonyms and make a wordle.

Students are creating a wordle of the main character in a book and the character traits, teachers are putting the "common word" in large print (x 4 or 5 entries) and then putting synonyms around to create visual thesaurus pages for students, students are copying and pasting their writing to make word clouds of the most used words, etc. The list goes on--educators LOVE Wordle!

Gift Giving

I created it as a gift to give to my boyfriend of his favourite song.

I give them as a birthday gift to the girls in our church. I ask parents and friends and other leaders for descriptions of the girls. Then I put them all together.

Wrote a love letter for my husband, Wordled it printed it on card stock paper and gave it to him for Valentine's Day instead of buying a store bought card. Then wrote the actual note behind it.

I created three of them. Two for my businesses and one for a gift (in a photo album) for my daughter when she had her baby.

In these scenarios, creating a Wordle is just the first step in a chain of events that may include the use of different media. These freeform answers are corroborated by the fact that 73% of respondents said they created their wordle to share with others.

4.2.4 (Mis)understanding Wordle

It is hard to gauge how well people understand a visualization technique. Because tag clouds have been widespread on the Web for years and because Wordle, like tag clouds, represents just a single variable (word frequency), it seems plausible that users understand how Wordle works. Nevertheless, one of the biggest surprises in this survey turns out to be people's incorrect perception of the Wordle technique. We asked three multiple-answer questions about different "dimensions" of Wordle:

1. On Wordle, what does the size of each word mean?
2. What does the direction (horizontal, vertical, diagonal) of each word mean?
3. What does the color of each word mean?

For each one of these questions, we provided the same set of answers, presented in random order:

- a) number of times a word is used
- b) emotional importance
- c) the meaning of the word
- d) nothing

Analysis of the answers revealed that comprehension interacts with age and gender. Table 4 shows a breakdown by gender and age of the users who did not know what font size stood for in a Wordle.

Table 4: Percentage of respondents who did not know what font size means in a Wordle

| | Male | Female |
|----------|------|--------|
| Under 20 | 35 | 49 |
| 20-30 | 12 | 18 |
| Above 30 | 19 | 31 |

Overall, younger people (under 20) tended to know the least about how each different dimension worked, with females doing worse than male users. One set of answers in particular stands out: word size. 35% of young males and 49% of young females did not understand the meaning of word size. Older females too (above 30) did not do so well: 31% did not understand what word size meant.

In addition, many people read meaning into the two non-coding dimensions of color and word direction. Of the two, color was more often interpreted as having meaning. To be sure, part of our results could

be due to the arguably leading nature of our questions. The very act of asking about the meaning of word color and direction might predispose participants into thinking that these dimensions should have meaning when they do not. But this would not explain the low comprehension numbers for the meaning of word size.

It is unclear whether this lack of understanding is specific to Wordle. How many of our respondents would give correct answers for a bar chart or scatterplot? Investigating lay understanding of charts would be an interesting area for future research.

If this confusion is specific to Wordle, it raises an important question. Given that the Wordle visualizations make little attempt to emphasize the visual encoding—no scale is shown on the charts, for instance—is it therefore deceptive? The design issues are complicated, since adding an ostentatiously scientific legend might interfere with some of the emotional uses we observed.

5 DISCUSSION

Given the prevalence of non-analytic Wordle usage, and the misinterpretation of Wordle's output, one might question whether it is being used as a "visualization" at all. Would it be more accurate to categorize these displays as decoration or entertainment? To answer this question, it's worth going back to the conventional definition of information visualization as "amplification of cognition by visual means" [1]. Cognition consists not only of scientific analysis but includes many other processes.

That visualization can aim at these other processes has been noted before, in the context of "casual infovis" [6]. In two respects, however, the term "casual infovis" fails to capture some of the most interesting uses of Wordle. First, not all the information viewed is personally meaningful: consider the case of students making wordles from literature in homework assignments. Second, the word "casual" itself seems inaccurate in describing many of the cases where people have used personally meaningful data. For the lovers who write valentine's messages, Wordle is, in fact, deeply engaging.

Terminology aside, it's worth pointing out two cognitive processes supported by Wordle that are not directly related to statistical analysis or insight: learning and memory. As discussed above, educators are a core group of Wordle users, and have found an eclectic set of uses for these word clouds. In some cases, teachers used Wordle to communicate ideas or concepts. Some educators used wordles as an entry point into discussion, asking students to explain the prominence of certain words, or to guess what a wordle of a text would look like. The engaging nature of wordles clearly was a factor in many of these scenarios.

It is typical to use visualization to illustrate and communicate. However, instead of simply presenting pre-made wordles to students, teachers often crafted exercises that involved the students making wordles themselves. In other words, the process of creation was a critical part of many of these classroom activities. The fact that students could take an active role clearly appealed to both teachers and students, and made Wordle into an effective learning tool.

Wordle also seems to be used to **augment memory**. An initially counterintuitive aspect of its usage is that many people visualized texts they were intimately familiar with. The woman who transformed her wedding vows, for instance, was unlikely to discover anything in the text she had not seen before. Instead, her self-described goal was to create a "memento"—that is, to help her remember a key life event. In our poll, 50% of users agreed with the statement that their Wordle "jogged their memory." It's worth noting how far this is from the usual framing of the value of visualization—in the context of discovery, finding out something you already know is considered worthless.

Previous investigations have noted the value of "reminding" in other contexts, however, especially social ones. The Contactmap system [15] was seen as helpful because it gave people awareness of ongoing relationships. Pousman et al [6] also pointed to social awareness as a benefit of non-analytic visualization. The strong emotional component to many of the "memento" wordles suggests an analogy with photography, as proposed in [11].

The ability of wordles to assist learning and memory seems directly related to their aesthetic qualities. Many people find Wordles to be objects of beauty and inspiration. The frequency of layout/font/color edits attests to the importance of typographic design for users. In comparing Wordles and tag clouds, for example users remarked on the superior coherence of Wordles: they offer visual composition where tag clouds present lines of text, they bring both unity and variety (form and counterform play off each other), something that is lacking in regular tag cloud design. In short, the visual depth of Wordles attracts the eye and keeps user attention for longer periods; a condition that is paramount to data analysis. Other visualization techniques might be also served by a design approach that emphasizes expressiveness as well as legibility.

5.1 Creativity and Participatory Culture

We argue that the feeling of creativity is central to the experience of using Wordle. Even the examples where Wordle aids learning and memory include elements of creation. For people making mementos, creativity is key to the experience; many people relate Wordle to scrapbooking. In the classroom, Wordle is not just a broadcast medium, but something that students can use themselves. One typically does not think of visualization as a creative outlet, any more than one would think of a microscope as an authoring tool. Rather than a scientific instrument, however, the type of visualization represented by Wordle may be more like a camera: a tool that can be used to document and create.

The activity surrounding Wordle seems to fit Jenkins' definition of participatory culture. The barriers to entry are low, and we see both self expression and, in the political Wordles, civic engagement. The system has technical and legal infrastructure for sharing creations. Finally, there is indeed informal mentorship. Consider the videos that people have made which describe Wordle; the lists of classroom uses; and the many blog comments full of wordle tips.

Viewed through the lens of Jenkins' work, some of the less obvious aspects of the Wordle phenomenon begin to make sense. The range of uses, from frivolous to serious, is characteristic of the participatory culture arc, with more whimsical uses leading to more sophisticated analysis—the element of “fun” attracts novice users of a system [5], and helps them learn how it works. Today a user may create a playful wordle of their favourite band, tomorrow they may point a treemap to the minutiae of government spending in their local community. Wordle and other playful visualization systems may serve as stepping stones for infovis literacy.

If it is true that Wordle has made the transition in people's minds from a viewing tool to an authoring tool, one might ask how it has done that. We suggest there are three ingredients: **user choice, artifact portability, and remixing power.**

User choice

Choice is inherent in the Wordle experience. Users can experiment with everything from color and layout to ways of sharing. The implications of such flexibility are twofold: each Wordle has the potential to look unique, and users are more likely to take ownership of their work. As our survey indicates, this ownership is a key element in how people relate to Wordle: by giving them choice, Wordle becomes an artifact of a participatory medium.

Artifact Portability

An authoring tool must be able to produce a lasting artifact. Several features of Wordle ensure that users' creations can persist. Not only can people create a web page with a persistent URL, as is typical of many online visualization sites, but it is easy to export to PDF. Beyond these technical capabilities, the legal framework of the Creative Commons license helps people distribute their creations. This infrastructure empowers users to think of Wordles as their property, inspiring further experimentation and creativity.

Remixing Power

Everyone has text they care about, whether emails, love letters, or speeches made by a hated politician. Text is almost never just “data.” Pointing Wordle to the latest cultural meme, be it a speech by the

president or the stimulus bill, and then sharing it, proves a quick and easy way of engaging in communal and civic meaning-making.

5.2 Implications for design

Our results suggest that users' attraction to Wordle rests upon two main points: design/typography choices and access to a creation tool.

It is hard to overstate the importance of design for Wordle's users. People are drawn to the interesting graphics and mention its striking nature as a key reason for playing with the tool. Not only do users want to visualize the gist of a text, they want a visualization with some punch. Unlike regular tag clouds with their tidy rows and tedious fonts, the dynamic and flexible graphics of Wordle provide the visual kick they seek.

Seemingly in opposition to conventional principles of visualization design—such as clarity and precision—Wordle's jumbled, non-linear nature resonates with the need on users' part to visualize, communicate and be expressive at the same time. It behooves visualization builders to explore the possibility of judiciously “bending the rules” of traditional design to allow for more expressive uses. For example, the choice to use a linear scale for word frequency in Wordle, as opposed to using a square-root method usually favoured by the infovis community, yields visually appealing results. Likewise, displaying text in different orientations might not help legibility but energizes the visual composition. If getting people excited about using visualization tools is one of the challenges builders face, good design choices should address that task. With its legions of users, Wordle suggests that people are more likely to spend time examining visuals they find appealing.

It would be easy to conclude that aesthetics alone are the reason that Wordle is popular. But to leave the results there would be missing an important point. If the Wordle experience contains an additional message for designers, it is that supporting the full creative cycle pays big dividends.

Wordle is one of the first opportunities non-expert users have of engaging with visualization production. The Wordle web site is designed to play a part in the online creative ecosystem. The software can easily consume textual artifacts, letting users remix information that is meaningful to them. Even more than earlier public visualization systems such as Many Eyes, the creation process itself is streamlined—no registration is required, giving instant results that encourage new users—yet allows people to vary enough parameters, to let them take ownership of the results. Finally, the site gives people the tools they need—such as explicit legal protection, persistent URLs, and high-resolution export capabilities—to take their Wordles elsewhere and continue the creative cycle. As visualization tools start to inhabit the Web 2.0 space, builders ought to support the exporting and embedding needs of users.

Finally, from a research perspective, Wordle proves that there is great value to visual experimentation on the open, public web. Even though the basic Wordle layout had been deployed internally in IBM since 2006 [8], we would not have thought to create a lab test to look at whether people felt “creative” after using a visualization—let alone whether it would help anyone get “boyfriend points.” Placing it on the web, and enabling anyone to use and share wordles, gave a unique window onto visualization usage.

6 CONCLUSION

The Wordle web site, though launched as a toy, has seen a wide diversity of applications. Major newspapers, teachers, and couples in love have all used Wordle for purposes ranging from the analytical to the emotional. Its remarkable level of popularity and the surprising number of usage scenarios led us to investigate wordle usage through a survey and examination of the internet record. **Our study revealed two main themes behind Wordle's broad uptake: the importance of design and the fact that the Wordle site works as a creation tool.**

Sophisticated use of typography and color lures users to the site and inspires them to create their own visualizations. In spontaneous blog statements and in our survey we found evidence that the visually striking nature of wordles drew users in and caused them to spend more time

examining the visualizations. The fact that people could customize their creations held additional allure, with our survey and internet study indicating that many users enjoyed adjusting parameters until their wordles looked how they wanted.

The ability for users to embed some of their own personality in their creations may account for the second theme we found: a vast number of people felt that making a wordle was a creative activity. We found many examples of people who treated wordle-based activities as akin to craft projects such as scrapbooking. User choice—ranging from aesthetic options to sharing alternatives—plays a key role in people’s creative process, allowing them to appropriate the artifacts they create on the Wordle site. Based on these findings, we propose that Jenkins’ notion of “participatory culture” is a useful framework for viewing Wordle, in that people actively create, remix, and share work. In particular, the ludic aspect of Wordle appeals to novice users, quickly introducing them to the creation process. In this sense, Wordles might function as a stepping stone for later, more complex use of visualization technologies.

At the same time, our survey revealed some potentially problematic aspects of the Wordle experience. A significant number of people do not understand the information encoding in Wordle. Our survey indicated strong age and gender differences in how wordles were interpreted, which suggests natural directions for future research. In addition to testing these findings in a lab setting, one might extend this investigation to how well the average person understands other very simple charts and graphs.

The main message, however, may be that viewing visualization as a creation or authoring tool suggests a set of new design points. Thinking of visualization users not as consumers of software but rather as a community of creators may help us move this technology into a vibrant future. In a world where visualization exists in the connected public web, user choice, artifact portability, and remixing power may be essential elements of success.

ACKNOWLEDGMENTS

We are grateful to Frank van Ham and Joan Dimicco for helpful suggestions, as well as the entire Collaborative User Experience group at IBM for many useful comments.

REFERENCES

- [1] Card, S., McKinlay, J., and Shneiderman, B., eds, (1999). Introduction, *Readings in Information Visualization: Using Vision to Think* Morgan Kaufmann.
- [2] Clark, J. *Word Hearts, from the Neoformix blog*, 2008. <http://www.neoformix.com/Projects/WordHearts/index.html>
- [3] Donath, J., Karahalios, K., Viegas, F. 1999. Visualizing Conversation. *HICSS-32*.
- [4] Hearst, M., & Rosner, D. Tag Clouds: Data Analysis Tool or Social Signaler? *In proc. of HICSS 41*, 2008.
- [5] Jenkins, Henry (2006). *Convergence Culture: Where Old and New Media Collide*. New York University Press
- [6] Pousman, Z., Stasko, J., and Mateas, M. Casual Information Visualization: Depictions of Data in Everyday Life. *IEEE Symposium on Information Visualization*, 2007.
- [7] Seifert, C., Kump, B., Kienreich, W., Granitzer, G., and Granitzer, M. 2008. On the Beauty and Usability of Tag Cloud, *12th Int'l Conf. on Information Visualization*.
- [8] Rivadeneira, A. W., Gruen, D. M., Muller, M. J., Millen, D. R. Getting our heads in the clouds: toward evaluation studies of tagclouds. *Proc. of CHI 2006*.
- [9] TagCrowd, <http://tagcrowd.com>
- [10] Tag Cloud Generator, <http://www.tagcloud-generator.com/>
- [11] Viégas, F.B., boyd, d., Nguyen, D., Potter, J. & Donath, J. Digital Artifacts for Remembering and Storytelling: PostHistory and Social Network Fragments. *In Proc. of HICSS-37*, 2004.
- [12] Viégas, F.B., Wattenberg, M., van Ham, F., Kriss, J., & McKeon, M. *Many Eyes: A Site for Visualization at Internet Scale*. Proc. of IEEE InfoVis 2007.
- [13] Viégas, F.B., Wattenberg, M., Tag Clouds and the Case for Vernacular Visualization, *ACM Interactions*, XV.4 - July/August, 2008
- [14] Wattenberg, M. Baby Names, Visualization, and Social Data Analysis. Proc. of InfoVis 2005.
- [15] Whittaker, S., Jones, Q., Nardi, B., Creech, M., Terveen, L., Isaacs, E., Hainsworth, J. ContactMap: Organizing communication in a social desktop. *ACM Trans. Comput.-Hum. Interact.*, Vol. 11, No. 4. (December 2004), pp. 445-471.

7 APPENDIX: WORDLE LAYOUT ALGORITHM

The most distinctive geometric aspect of a Wordle is the layout algorithm, which packs words to make efficient use of space. While many space-filling visualizations exist, they typically work by recursively subdividing regions. The Wordle algorithm, by contrast, is agglomerative and builds up the final layout by adding one word at a time.

The input to the wordle algorithm is a list of words, each of which has a numeric “weight.” For each word, Wordle constructs a font with a point size proportional to the word’s weight, with the highest-weight word assigned an arbitrary font size of 1,000. The algorithm then uses that font to generate a shape, expressed as a collection of spline paths. At this point, the goal is to position these shapes in a non-overlapping manner so that (a) space is used as efficiently as possible; (b) each word is close to an “ideal” position as defined by a placement criterion.

The overall strategy is a randomized greedy algorithm in which words are given a position in space one at a time. Wordle offers the user a choice of placement: *center-line* and *alphabetical center-line*. Both strategies place words near the horizontal center-line of the playing field (not necessarily directly on it, but scattered away from it by a random distribution). The alphabetical strategy additionally sorts the words alphabetically, then distributes their preferred *x*-coordinates across the playing field.

Layout proceeds according to this pseudocode:

```
sort words by weight, decreasing
for each word w:
  w.position := makeInitialPosition(w);
  while w intersects other words:
    updatePosition(w);
```

The two key procedures here are “makeInitialPosition” and “updatePosition.” The makeInitialPosition routine picks a point at random according to a distribution that takes into account the desired overall shape, and, if desired, alphabetical order. The updatePosition routine moves the word on a spiral of increasing radius, radiating from the word’s starting position.

The updatePosition routine is aware of constraints on the overall shape of the Wordle. Constraining the layout to a rectangular shape causes updatePosition to prefer positions inside of the strict boundaries of the playing field; a blobby overall shape accepts boundary violations. The rectangular constraint is relaxed when the spiral radius exceeds either playing field dimension.

7.1.1 Performance Tuning

A naïve implementation turns out to be very slow. The bottleneck is the tests for intersection between the word to be placed and the words that have already been placed. Determining whether spline-based shapes intersect is expensive; and even if the first position attempted for each word turned out to be acceptable, the total number of collision tests for a layout would grow quadratically in the number of words. To help performance, therefore, we use several tactics.

First, we create a “hierarchical boundig box” for each shape. Creating this data structure takes some time up front but speeds the overall process.

Second, we create a spatial index of words already placed—a quadtree structure that lets us efficiently cull shapes from the list of comparisons for a particular word.

A final improvement stems from the observation that if a word A intersects some other word B, then it’s very likely that A will still intersect B after moving A slightly. This allows for a significant speed-up in the position-updating routine because the program can cache the candidate word’s most recently intersected word, and test it first at the new position.