

Embroidered Ephemera: Crafting Qualitative Data Physicalization Designs from Twitter Data

Anne Sullivan

StoryCraft Lab
Georgia Institute of Technology
Atlanta, GA 30332 USA
unicorn@gatech.edu

Abstract

Embroidered Ephemera explores the juxtaposition of tweets—which can take as little as a few seconds to create, share, and ultimately disappear—and embroidery, which can take hundreds of hours to create yet also last hundreds of years. Embroidered Ephemera is an online tool that allows users to input a Twitter user or hashtag and then generates embroidery sampler designs based on collected tweets. The generator uses an assortment of text analysis tools to choose words, motifs, and colors based on the chosen tweet from the specified user or hashtag. The text analysis influences the possibility space for the aesthetic aspects of the generated sampler design but is not used to define the output such as would be seen in typical quantitative data visualization.

Introduction

Embroidery is a form of handcraft that has been around since at least 30,000 BC – which is the date of the earliest known fossilized remains of an embroidered piece – and is a craft practice that continues through today. Embroidery uses a handheld needle to apply thread or yarn as a decorative feature to a piece of cloth or other material. Through the years, embroidery has been used to embellish clothing and other items, and due at least partially to the amount of time that is required to create the needlework patterns, it imbues a sense of importance and value to whatever item has been embroidered.

Embroidery encompasses a large variety of stitch types, each requiring practice to develop the skills required to create high quality designs on valued clothing or items. To gain proficiency, different stitches would historically be practiced on scraps of cloth. These scraps also became a way to share stitch knowledge between people and allowed for easier information exchange. Over time, this practice evolved into more aesthetically driven and decorative designs. These pieces are now referred to as embroidery samplers, and while partly ornamental, they are still used as a way to show expertise at one or more stitch types.

As samplers have evolved, they began to become more complex, and often include a combination of motifs,



Figure 1: A cross-stitch embroidery sampler made by Roxana Corson, age 9, ca 1826.

alphabets, quotes, figures, borders, etc. These samplers take weeks or months to complete, and in some cases last for centuries (see Figure 1).

These designs also began to often include the name of the person who embroidered it and the date that it was completed. This is not typical of other embroidered designs, and it is likely this practice developed as samplers began to be known as a “specimen of phenomenal achievement” (Huish, 1913) and displayed as their own stand-alone art piece. Because of the length of time embroidery can survive and because samplers regularly include names and dates; it has provided a legacy for the embroiderers; both highlighting



Figure 2: A design generated by Embroidered Ephemera using tweet data by influential Twitter user @dril.

their skills and what things they valued enough to include within the sampler.

In contrast to the time commitment required for embroidery, Twitter is an online format that allows users to quickly share their thoughts to a global audience, in a matter of seconds or minutes. In turn, a tweet can then easily be lost in a sea of other tweets within a very short time frame. Because of this, while Twitter has become a powerful cultural force, the vast majority of tweets, when taken on their own, are not considered particularly valuable by the author or audience.

Embroidered Ephemera is a casual creator that was designed to explore this juxtaposition between cross-stitch embroidery samplers and tweets. What could it look like to have tweets highlighted as something valuable and special like embroidery? In what ways could we capture some qualities of the tweet author and what is valuable to them?

To answer these questions, we created an online tool that generates sampler designs based on data derived from tweets from a user-specified person or hashtag. The system combines these tweets with generated motifs and crowd-sourced color palettes to constrain the possibility space of the generator which will then generate an embroidery sampler design (see Figure 2).

In the rest of this paper we will discuss the Embroidered Ephemera system in more detail and describe the design decisions that were made while creating it. We will also describe what we learned by creating this system and discuss possible future work.

Related Work

Embroidered Ephemera draws inspiration from many fields including craft research, graphic design, steganography, craftivism, casual creators, qualitative data visualization, feminist data visualization, and data physicalization.

In particular, the rich history of encoding data in textile crafts was a strong influence for this project. Textiles have a long history of use in steganography, most famously fictionalized by Charles Dickens in *Tale of Two Cities*, with the character Madame Defarge, who knit names of nobles to be sent next to the guillotine (Dickens, 1899). This was not just fantasy, as there are records of women knitting information about train schedules during wartime in 1914 Belgium (Witkowski, 2015).

More recently, textile creators have been exploring aspects of creative and personal expression with the use of covertly embedding data into textile creation. For example, Spyn (Rosner, 2008) uses infrared ink to allow the knitted artifact to be “read” through computer vision to recall what was happening to the knitter during the time they were knitting that area. Similarly, Haring’s “Subtle Distress” (Yidan, 2019) uses purl and knit stitches to represent Morse code within the stitches of a sweater.

Most steganography looks to camouflage data while still being decipherable, however with Embroidered Ephemera we are not aiming to hide information, but rather display the data beyond a purely informational way. This aligns with the ideals of feminist data visualization, particularly embracing affective experiences of the viewer (D’Ignazio & Klein, 2016).

This is often seen with textile-based data physicalization, which gives people a chance to interact with the data in a tangible and more affective way (Jansen et al., 2015). For instance, many crafters have knit, crocheted, or sewed blankets or quilts that represent local or historical temperature data (Onion, 2020). The final artifacts may encode data, but they can also be appreciated as both a useful object and as an aesthetic object. Most of these projects use quantitative data, but for Embroidered Ephemera, we were interested in using qualitative data for designs that could be used for data physicalization.

It was also important to us that the tool was lightweight, easy to use, and playful. For this, casual creators provided the inspiration as a “pleasurable exploration of possibility space” (K. Compton & Mateas, 2015). Smaller casual creators such as the magic shop generator and other RPG-based tools by donjon (*Donjon; RPG Tools*, n.d.) with very limited controls encourages more exploration. However, unlike these simpler generators, we wanted Embroidered Ephemera to use data to shape the generative space of the system.

System Details

Embroidered Ephemera is a web-based tool that generates embroidery sampler designs that can be used to create a physically embroidered piece. The system is written in JavaScript using jQuery and JSON and is available open source at <https://github.com/anneandkita/twittersampler>.

The focus of this project is to generate designs that are informed by the traditions of embroidery samplers but use modern-day data as inputs for the generator. The goal was to create samplers that could aesthetically blend with traditional embroidered samplers. Therefore, aesthetics is a large consideration for the output of the system. The output is not meant to be clearly interpreted data visualization; instead the design is aesthetically influenced by the data to create an affective experience.

As embroidered samplers have evolved over centuries, there have been a number of sampler styles. For this project, we chose to use the “band sampler” style, which is made up of rows of horizontally repeating motifs, and often include text. Band samplers became popular when the price of fabric made it difficult to get large pieces of cloth, as the design can easily be modified to fit different aspect ratios and sizes. This was chosen as a well-suited style to represent the limited character size of tweets, which often causes users to similarly modify their words to fit within the limit.

As mentioned earlier in the paper, embroidery encompasses a large number of stitch types. For this project, we chose to limit it to cross-stitch, which is a stitch type that is made by creating an X-shaped stitch over the weave of the cloth. Cross-stitch is used regularly in samplers as it can be used for many different types of designs, and each X can be viewed as a pixel, which is well-suited for our computational needs.

Output Generation

To start using Embroidered Ephemera, an interactor enters a Twitter username or hashtag onto the site. From there, the system uses the Twitter API to search the last 30 days of tweets from that user or hashtag (*Docs*, n.d.). The tweets are ranked based on engagement – calculated based on the number of comments, re-tweets, and likes – with higher engagement giving the tweet a higher rank. Using the ranks as weights, the system makes a weighted random choice, with higher engagement tweets more likely to be chosen.

The text of the tweet is converted into a cross-stitch font using a custom-designed library of font glyph patterns. Once converted to a cross-stitch pattern, the height of the text block is calculated and placed using the rule of thirds as a rough guideline (Peterson, 2015). The text block is always placed higher than the bottom one third of the design, to give the overall composition a balanced feel (*ibid*). Additionally, the name of the twitter user and the date are converted to the cross-stitch font and placed within the design to mimic the dated signature of the sampler creator. The block containing the user and date is placed at approximately the top or bottom one-third line, opposite of the text block to create balance and symmetry.

From here, the system generates the rest of the design around the text blocks. Since Embroidered Ephemera generates band samplers, the system generates one band (or row) at a time. A row can either be a repeating graphic that is tiled horizontally, or a set of words that is taken from a list of most-used words.

Most used words

The sampler design has the possibility of including one or more bands made up of words. This is inspired by traditional samplers which would sometimes include important concepts to the creator such as home, love, or family. To find what words were important to the twitter user or hashtag specified, the system uses the 30 days of tweets returned initially by the Twitter API to find the top 10 most used words, modified by uniqueness.

To accomplish this, we use term frequency–inverse document frequency (TF-IDF) analysis. The analysis calculates term frequency using the 30 days of tweets and makes use of a corpus of tens of thousands of tweets (primarily written in English) as the document for the inverse document frequency. This means that words such as “the”, “and”, “you”, etc. will be less likely to appear.

The list of top 10 unique words can then be used within the design of the sampler. Once a word has been used in a band in the sampler design, it cannot be used again within that same design. Once all the words have been used in a design, no more bands with words can be created. Bands with multiple words use the “+” character between words to help keep the word lists from appearing to be part of the original tweet.

Graphics

Once a tweet has been selected, the system uses Google’s sentiment analysis API (*Analyzing Sentiment | Cloud Natural Language API*, n.d.) which calculates a numerical “positivity” rating and a magnitude (or confidence in that rating) based on the chosen tweet. From these two numbers, the system calculates an “activation” rating with lower positivity scores having lower activation. The activation rating is put into one of three categories: low, medium, or high, with lower numbers being put into the low category, etc. The magnitude returned by the sentiment analysis API is used as a modifier such that lower magnitudes (lower confidence) trend towards medium activation. The activation rating is then used as a constraint for generating graphics and colors within the sampler design.

For the graphics, we created a library of ~100 horizontally tiling cross-stitch designs. These designs are either created specifically for this project or are sourced from open source and public domain patterns, and feature both traditional and modern motifs. All designs were converted for our library to work with our system.

Each design in the library is hand-tagged with an activation rating based on the design principle of movement, which relates to the way the viewers eye travels through a piece of art (Peterson, 2015). Therefore, motifs with squares, grid-like qualities, or primarily horizontal or vertical lines are rated low activation, organic motifs or curvy, flowy designs are rated medium activation, and angled or jagged motifs are rated high activation (see Figure 3).

To choose a row, the system uses the activation rating of the tweet and chooses randomly from the designs that have



Figure 3: The tweet on the left was designated as a “medium” activation tweet, therefore the repeated motifs include a lot of organic and curved lines. The tweet on the right on the other hand was categorized as “low” activation, and therefore uses squares, grid patterns, and horizontal or vertical lines in the design. While the colors are similar, the difference in motifs make the two pieces noticeably different.

the same activation rating, with a small chance of choosing a motif that has an activation rating one step away from the tweet’s activation rating (e.g. a low activation tweet has a small chance to include a medium activation motif, but would never have a high activation motif.)

Colors used

In every design, black is used as the primary color, which is used for all text and is included as a color in almost every graphic motif. However, motifs can have up to two other colors, so other colors need to be generated for each sampler design.

To choose colors, the system again uses the activation rating of the tweet. Here, we use US-specific culture-based views of colors to assign a range of colors to each activation level. For instance, blue is considered a calm color in the US, so it is associated with low activation. For our system we use the following list:

- Violet or blue – low activation
- Green or yellow – medium activation
- Orange or red – high activation

One of the two colors is chosen randomly from the associated activation rating and is used in combination with the COLOURlovers API (COLOURlovers API Documentation, n.d.). The system requests a random palette from the top 100 rated palettes on COLOURlovers which includes the specified color. From the palette that is returned, the perceived brightness is calculated for each color. Any color that has a

perceived brightness too close to the default background color is removed from the list to ensure the motifs will show up against the background.

Two colors are chosen from the remaining colors in the palette, and black is used if there are not enough colors remaining. As the color palettes are rarely monochromatic (for instance a palette with green in it could also include pink and blue), and the colors are chosen randomly, the initial seed color does not always appear in the final design. We chose to allow this to increase the generative space for the colors, as always using a specific color led to an overly limited aesthetics.

Discussion and Future Work

As detailed earlier, Embroidered Ephemera was designed to explore the contrast between the speed and ephemeral nature of Twitter versus the meditative and lengthy practice of embroidery. There is an immediacy to posting to Twitter that in many ways is similar to that of a generator such as Embroidered Ephemera. This speed of creation often leads to a very “noisy” space, and it is difficult for something to stand out.

Towards this point, the average typing speed on a phone keyboard is ~25 wpm. This means that someone can write and send a full-length tweet in a little less than 2 minutes, and this can be done even faster when using a keyboard or sending a shorter tweet. Embroidered Ephemera can create

a new sampler design in approximately 4-5 seconds. This means that it would take 6-8 minutes to create 100 designs.

In contrast, embroidery is a much slower process. Even with cross-stitch which does not have the cognitive load of switching stitch types, it is a labor-intensive process. To embroider one of the generated samplers from Embroidered Ephemera on 18-count Aida cloth (a type of woven cloth that has 18 holes per inch), it takes approximately 1 hour to stitch one line of one color. It is important to note that this refers to one line of “pixels” or stitches, not one full design band (see Figure 4). At this rate, based on the size of the sampler designs, it would take roughly 300 hours to complete. Therefore, in the time it takes to create one sampler, someone could generate ~28,000 designs, or send ~9,000 full length tweets from their phone.

With such a large time investment, the user experience of the system becomes particularly important.

User Experience

While Embroidered Ephemera is in early stages and therefore has not had extensive user testing, we have gathered informal feedback as well as noted our personal experiences while using the system.

There are two stages of the experience with the system. The first is using the generator, while the second is creating the embroidery from the generated design. While a number of users have interacted with the generative system, only the author has taken the next step in working on an embroidery of a generated design.

This mismatch in engagement likely speaks to the unequal human labor requirements between generating a design and hand embroidering that design. The speed at which the user can generate new designs, and the level of input the user has with the system fits the diversionary aspect of a casual creator. However, while embroidery and other craft-based hobbies are also sometimes seen as diversions (due to the devaluation of reproductive labor; crafting is not economically viable, therefore it must not be a serious undertaking (Hughes, 2012)), the time investment and depth of the interaction show a level of commitment to the process that is not found in casual creators.

The difference in commitment also touches on another challenge in generative design; that of perceived value of the output of the system by the user. Even when there is initial perceptible uniqueness (K. E. Compton, 2019) in the output, as the user explores the generative space, that uniqueness diminishes the more designs are generated. This is even more pronounced with a shallow generative space such as is found in Embroidered Ephemera.

Beyond this, in textile crafting (and likely most crafting), creative expression is one of the core values of practitioners (Sullivan et al., 2018). However, with the current version of Embroidered Ephemera, the generative system is doing all of the design work: choosing a tweet, choosing borders, choosing colors, choosing a layout. This leaves the user with very limited opportunities for expressing their own creativity in the design generation beyond choosing a user or hash tag and hitting the re-roll button.

Our experience creating the embroidered artifact aligns with these observations. Because all of the creative decisions have already been made, the act of creating the embroidery becomes machine-like and thoughtless. And while the opposing values of tweets and embroidered samplers leads to some playful interactions and design, this is shallow enough that it is unlikely to carry the embroiderer through the hundreds of hours required to finish the physical artifact. In its current incarnation, the user’s creative expression only exists where the embroiderer chooses *not* to follow the design.

This could be addressed by allowing the user to retain more creative control during the design process and/or to encourage moments of user creativity during creation of the artifact. To work within the system as it exists now, creative expression exists primarily at the curation level; what Twitter users, tweets, and designs are chosen to be embroidered.

The curation of tweets as creative expression is what we relied on in our experience of choosing tweets to embroider. In the example seen in Figure 4, a selection of tweets by political figures were chosen to highlight the differences in tone and content and how this might affect the final design. The tweet shown was chosen in part because it characterizes the general tone of the author’s tweets and highlights the contrast between this style of delivery versus that of Michelle Obama as seen in Figure 3. It was also chosen because it brings into focus the conflict of politics and twitter; twitter is designed for short form and quick discourse but is being used by the current US administration for political statements. By committing a quick and reactionary tweet to something that as long-lasting as embroidery, it highlights the power and control those words have, even if they are not intended or used that way.

Privacy

An issue we had not considered when starting this project is that of data privacy. While it is possible to search any username on twitter and see their tweets if their privacy settings are set up to do so, our casual creator is still bringing focus to tweets that may not be something that the original author may want. Because the Twitter API and our tool do not ask for user permission, it raises ethical concerns.

In particular, a design can be saved and shared which contains a tweet which the original author may have deleted or may not have wanted to give that type of attention to. Given the length of time that embroidery can survive, if a design is embroidered into a physical artifact, this means that the tweet may outlive the author themselves.

While this is beyond the scope of this paper and an area of active research, we do not have a current solution to this beyond asking people to be considerate when using the tool, or not allowing other people to use it at all. In the future if and when the Twitter API is extended, we would like to change the tool such that people can only look up their own username or those of public figures.

US-Centric

Finally, a large constraint on our system is that it is no better than a random generator when using tweets that are not in



Figure 4: Clockwise from top left: The original tweet. The design created by Embroidered Ephemera based on the original tweet. The beginning of a physical embroidered piece based on the design.

English or when looking beyond color-based cultural references from the US. In the future, this could be addressed by using either the location of the tweet author or language of the tweet to modify how the data is interpreted. This would also rely on the sentiment analysis API having support for different languages beyond English.

Future Work

In many ways this is a data visualization project, however the end goal was never to create an artifact that affords the data to be easily interpreted such as is expected from traditional data visualization. Instead, the data is analyzed and used to shape the aesthetics of the piece, but it does not define them entirely. We made this choice in part due to the fact that we are using qualitative rather than quantitative data.

The system currently quantizes the sentiment analysis results by assigning them to the three activation levels. This leads to a lack of nuance within and between the designs, although that lack of nuance may help with perceptible uniqueness and differentiation between categories of designs.

We would like to further develop and incorporate the activation concept by using continuous values and refining the way we categorize colors and graphical motifs. It is also possible to incorporate this concept into more of the design by looking at heights of motifs, different types of samplers beyond the band sampler, and incorporating colors in the text and background. The system would also be improved with additional motifs and fonts in the library. All of these changes would increase the generative space and allow the

user to explore more areas of the possibility space through the generator.

Finally, the user's exploration of that space is currently done through a very limited user interface. While this does invite exploration, the interaction is very shallow, and limits the user's creative expression as mentioned above. Adding the ability to modify meaningful pieces of the design would provide a better balance between control and ease of use. For instance, being able to lock or re-generate a specific tweet, color palette, or row design would give the user ways to further influence the design while still allowing room for exploration.

Conclusion

In this paper we have presented Embroidered Ephemera, a casual creator which designs traditional-inspired embroidery samplers from Twitter data. Text analysis is used to shape the aesthetics of the design, by constraining the options of motifs and colors available to the generator. Each sampler design is therefore aesthetically influenced by the data but is not defined by it.

References

- Analyzing Sentiment | Cloud Natural Language API.* (n.d.). Google Cloud. Retrieved June 2, 2020, from <https://cloud.google.com/natural-language/docs/analyzing-sentiment>
- COLOURlovers API Documentation.* (n.d.). Retrieved June 2, 2020, from <https://www.colourlovers.com/api>

Compton, K. E. (2019). *Casual Creators: Defining a Genre of Autotelic Creativity Support Systems* [UC Santa Cruz]. <https://escholarship.org/uc/item/4kg8g9gd>

Compton, K., & Mateas, M. (2015). Casual Creators. *ICCC*, 228–235.

Dickens, C. (1899). *A tale of two cities*. C. Scribner's sons.

D'Ignazio, C., & Klein, L. F. (2016). Feminist data visualization. *Workshop on Visualization for the Digital Humanities (VIS4DH)*, Baltimore. IEEE.

Docs. (n.d.). Retrieved June 2, 2020, from <https://developer.twitter.com/en/docs>

Donjon; RPG Tools. (n.d.). Retrieved June 2, 2020, from <https://donjon.bin.sh/>

Hughes, C. (2012). Gender, craft labour and the creative sector. *International Journal of Cultural Policy*, 18(4), 439–454. <https://doi.org/10.1080/10286632.2011.592187>

Huish, M. B. (1913). *Samplers & tapestry embroideries*. London, New York [etc.] Longmans, Green and co. <http://archive.org/details/samplerstapestry00huisrich>

Jansen, Y., Dragicevic, P., Isenberg, P., Alexander, J., Kar-nik, A., Kildal, J., Subramanian, S., & Hornbæk, K. (2015). Opportunities and Challenges for Data Physicalization. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3227–3236. <https://doi.org/10.1145/2702123.2702180>

Onion, R. (2020, February 8). *The Quilters and Knitters Who Are Mapping Climate Change*. Slate Magazine. <https://slate.com/technology/2020/02/quilts-knitting-cross-stitch-climate-change.html>

Peterson, B. (2015). *Learning to see creatively: Design, color, and composition in photography*. Amphoto Books.

Rosner, D. K. (2008). *Spyn: Weaving Stories into Hand-crafted Artifacts*. 19.

Sullivan, A., Salter, A., & Smith, G. (2018). Games crafters play. *Proceedings of the 13th International Conference on the Foundations of Digital Games*, 1–9. <https://doi.org/10.1145/3235765.3235802>

Witkowski, J. (2015). Knit for Defense, Purl to Control. *In-Visible Culture*, 22. <https://ivc.lib.rochester.edu/knit-for-defense-purl-to-control/#fn-3529-26>

Yidan. (2019, April). *Hiding in the Light*. <http://textileartscenter.com/blog/hiding-in-the-light/>