CIS 496 / EAS 499 Senior Project
Project Proposal Specification
Instructors: Norman I. Badler and Joseph T. Kider. Jr.

Caplet: An Interactive iPhone and Web Application for Personal Data Collection, Visualization, and Analysis

Brynn Shepherd & Nirav Sanghani

Advisor: David Comberg (comberg@design.upenn.edu)

University of Pennsylvania

ABSTRACT

*Our project enables individuals to both collect and analyze their personal data. We will design, build, and test an iPhone and web application that will allow users to collect **ONE** unit of personal information a day, which we are calling a "caplet." Each one of these will be a time capsule of data, capturing a particular moment in a user's life, and including at least the following dimensions: image, text, location, time, mood and weather. While the iPhone application will function as a tool for our users to enter their content, the accompanying web application will provide them with a means of browsing and visualizing it. The fact that users can only post one caplet a day will cause them to think more carefully about what they choose to record, and will therefore make that information more meaningful.*

While various web services already exist to facilitate data collection and analysis, few attempt to achieve both simultaneously, and few take full advantage of a mobile platform. In addition, most of these services allow users to post an unlimited amount of information each day, cluttering the Web with lower-quality content and making it increasingly difficult to filter out relevant information from noise. Our project does not attempt to completely resolve this problem, but rather presents an alternative to content creation for the Web and a more effective approach to structuring information.

Project Blog: http://acapletaday.blogspot.com

1. INTRODUCTION

In today's world, we are often overwhelmed by the quantity of information we encounter each day. The Internet and other modern computing technologies have made it increasingly simple for anyone anywhere to create content and share it with others. This has resulted in a proliferation of information, both digital and non-digital. We see this trend has having both positive and negative consequences; on the one hand, new technologies have democratized the production of content and made it easier for less powerful voices to be heard. On the other, however, we believe it has become increasingly difficult for people to organize and interpret the information that they are presented with, simply because there is so much of it and it often arrives in an unstructured form.

While computers are partly responsible for this problem of "information overload," they can also help us construct a solution. Scripts and programs allow us to structure, organize, and process information more powerfully than we've ever been able to before. Combined with effective user interfaces, they can make data collection and analysis accessible to interested people would would otherwise be overwhelmed. While various services already exist to facilitate data collection (Twitter, Foursquare) and analysis (ManyEyes, Daytum), few attempt to achieve both.

Our proposed solution is an integrated system through which individuals can both collect *and* analyze their data. We plan to design, build, and test an iPhone and web application that will allow users to collect **ONE** unit of personal information a day, which we are calling a "caplet." Each one of these will be a time capsule of data, capturing a particular moment in a user's life, and including at least the following dimensions: image, text, location, time, mood and weather. While the iPhone application will function as a tool for our users to enter their content, the accompanying web application will provide them with a means of browsing and visualizing it. The fact that they can only post one caplet a day will cause them to think more carefully about what they choose to record, and will therefore make that information more meaningful.

In the words of information design's pioneer, Edward Tufte, "Clutter is a failure of design, not an attribute of information." [T06] With this in mind, we seek to design our application in such a way that it transforms raw data into meaning.

1.1 Design Goals

Our project targets people who see value in recording data about themselves in a structured way and are already accustomed to doing so on a daily basis using their phones. Our primary goal in designing Caplet will be to simplify its user interfaces (both for iPhone and web) as much as possible without compromising its underlying functionality and the richness of the data it collects. The experience of using the application will be intuitive and unintrusive, making the process of posting a caplet fun rather than a chore.

1.2 Projects Features and Functionality

Caplet has two main components: an iPhone application that allows users to collect data and a web application that allows them to analyze it. The data will be structured in the form of daily caplets, multivariate pieces of information that include a photograph taken by the user, text composed by the user, the user's current location, the current time, the user's mood, and the current weather. The application will only allow a user to post a single caplet per day.

With its portability and powerful functionality, we feel that the iPhone is an ideal platform for on-the-go data collection. Many of the components can be obtained "for free" from the user's perspective; location, for example, can be computed automatically by the iPhone's built-in GPS, while time can be provided by its internal clock. We can even use the GPS data to compute the user's zip code, which we can then correlate with the current weather at that location.

Once a significant amount of data has been collected, we will try to uncover interesting patterns and trends by visualizing it on the accompanying website, probably by using a special-purpose technology such as Processing. We might, for example, compute a graph of a user's location over time, and see how that compares to their mood. Because the data will be collected in a structured format, it will be easier for us to extract valuable insights from it and display those insights back to our users.

2. RELATED WORK

As mentioned in the introduction, many services exist for personal data collection: Facebook allows users to upload photos and status updates; Twitter allows users to broadcast messages of under 140 characters; Foursquare and Gowalla allow users to "check-in" at their current location using the iPhone's built-in GPS. These last two location-focused applications are probably the most similar to Caplet in terms of their integration of an iPhone app for data entry and a website for data display. None of these services, however, provide a good means of visualizing or analyzing the data that has been collected. In addition, since the number of posts or check-ins per day is unlimited, the data is often noisy, uninteresting, and overwhelming in quantity. If users were only allowed to post one status update on Facebook per day, it would probably not be a quiz result.

There are also a handful of existing services for data visualization and analysis. ManyEyes, Daytum, and Processing [F08] and [FR08] do a particularly good job of visualizing information, but the cost of data entry from the user's perspective remains high. We feel that this cost can be greatly reduced by transferring the "data-in" interface from the web to a mobile device.

How to effectively design information is a hot topic right now in various design and technology circles, and we have been greatly inspired by the work of the following groups and individuals (see also [BW08] for Lee Byron's paper on graph visualizations):

- Jonathan Harris: We Feel Fine, The Whale Hunt, 10x10
- Ben Fry: Zipdecode, On the Origin of Species, Piet Mondrian Goes to the Super Bowl
- Lee Byron: Stacked Graphs
- Aaron Koblin: Flight Patterns
- Nicolas Feltron: 2008 Annual Report

The New York Times Graphics Department: A Peek into Netflix Queues, How Different Groups Spend Their Time, Box Office Revenues, Olympic Medal Count

3. PROJECT PROPOSAL

We intend to create companion iPhone and web applications that allow users to gather and visualize personal data as described above.

3.1 Anticipated Approach

The Caplet iPhone application will first allow the user to sign up for an account or link their app with an existing account (either created on the Caplet website or from a third-party authentication source such as Facebook Connect). From that point on, the iPhone app will expose only the functionality needed to compose that day's caplet. While posting a caplet, the user will be able to enter a limited amount of text into a field (exact number of characters to be determined), upload an image using iPhone's photo framework, and select a mood from a pre-determined list of options (e.g., "happy," "sad," "excited," etc).

Meanwhile, the phone's location and the current time will be determined independently by the app when the user submits the data. Once the user does so, he or she may only edit and resubmit the *text* portion from that day (in case of typos, for example). After the day is over, the post will be "buried" like a time capsule and locked from editing.

Upon receiving a post, the server will use the location and time to look up the weather that the user was experiencing upon submission and, along with all the rest of the data, will store it in our database.

When a user visits the website, he or she can log in, but cannot enter data. The primary purpose of the website is to browse previous caplets and to use the visualization tools to analyze the data stored in the database. There will also be a way for a user to export his or her data in a structured format.

3.2 Target Platforms

We will build the native iPhone application using Objective-C and the Cocoa framework, and the web back-end using the following server stack: Python running on the Django framework using the MySQL DBMS for storage and persistence. Finally, the website's browser-facing front-end will be built using HTML, CSS, and JavaScript running on the jQuery framework. The visualizations will be generated using these as well as Processing.

3.3 Evaluation Criteria

We plan to conduct a controlled user study to test whether users are more likely to collect data with our iPhone application as opposed to existing technologies such as web-based data entry. We will do this by assigning different sets of users the same task, but asking them to complete it in different ways. One group, for example, will be asked to track their daily location for a week using our iPhone application, while another will be asked to do so using Twitter or an even simpler web interface. If the group using our application accomplishes the task more consistently, we will have proved the value of our mobile approach to personal data collection.

We also will use our visualizations and analyses to ensure that our data is interesting and not simply noise. For example, text entry can be analyzed by word frequency, locations can be plotted on a map, and moods can be graphed on a chart. We plan to evaluate the data by the degree to which we are able to extract interesting patterns and trends from it.

4. RESEARCH TIMELINE

Realizing that a significant amount of data will need to be collected before we can do anything useful with the visualization and analysis component of our project, we plan to manually populate our database with sample data over the course of the first few weeks. Once the basic functionality of the iPhone application is complete, we will install it on a handful of our friends' phones and encourage them to start posting daily caplets as well. This should provide us with a sufficiently large data set needed to start working on the web application. See Figure 1.

Project Milestone Report (Alpha Version)

- Background reading complete (see references).
- Database design and implementation complete.
- Database populated with sample data and extensive test data collection in progress.
- iPhone application development in progress: basic functionality complete.
- Website application development in progress: basic functionality complete.

Project Final Deliverables

- iPhone application fully functional.
- Website fully functional.
- iPhone app UI designed, polished and user-tested.
- Website UI designed, polished and user-tested.
- Database extensively populated with real-world data.
- Visualizations built and integrated into website.

Project Future Tasks

- Visualization API
- Data API
- Android/Palm Pre applications

APPENDIX

A. Links

a. Twitter: http://twitter.com/

b. Foursquare: http://foursquare.com/

c. ManyEyes: http://manyeyes.alphaworks.ibm.com/manyeyes/

d. Daytum: http://daytum.com/

e. Processing: http://processing.org/

f. Facebook: http://facebook.com/

g. Gowalla: http://gowalla.com/

h. Jonathan Harris: http://www.number27.org/

i. We Feel Fine: http://wefeelfine.org

j. The Whale Hunt: http://thewhalehunt.org/

k. 10x10: http://tenbyten.org/

l. Ben Fry: http://benfry.com/

m. Zipdecode: http://benfry.com/zipdecode/

n. On the Origin of Species: http://benfry.com/traces/

o. Piet Mondrian Goes to the Super Bowl: http://benfry.com/writing/archives/264

p. Lee Byron: http://www.leebyron.com/

q. Stacked Graphs: http://www.leebyron.com/else/sreamgraph

r. Aaron Koblin: http://www.aaronkoblin.com/

s. Flight Patterns: http://www.aaronkoblin.com/work/flightpatterns/

t. Nicolas Feltron: http://feltron.com/

u. 2008 Annual Report: http://feltron.com/index.php?/content/2008_annual_report/

v. New York Times: http://nytimes.com/

w. A Peek Into Netflix Queues: http://www.nytimes.com/interactive/2010/01/10/nyregion/20100110-netflix-map.html

x. How Different Groups Spend Their Time: http://www.nytimes.com/interactive/2009/07/31/business/20080801-metrics-graphic.html?ref=multimedia

y. Box Office Revenues: http://www.nytimes.com/interactive/2008

/02/23/movies/20080223_REVENUE_GRAPHIC.html

z. Olympic Medal Count: http://www.nytimes.com/interactive/2008/08/04/sports/olympics/20080804_MEDALCOUNT_MAP.html

References

[T01]	Tufte, E. *The Visual Display of Quantitative Information.* Graphics Press, 2001.
[T97]	Tufte, E. *Visual Explanations.* Graphics Press, 1997.
[T90]	Tufte, E. *Envisioning Information.* Graphics Press, 1990.
[T06]	Tufte, E. *Beautiful Evidence.* Graphics Press, 2006.
[F08]	Fry, B. *Visualizing Data.* O'Reilly Media, 2008.
[FR08]	Fry, B., Reas, C. *Processing: A Programming Handbook for Visual Designers and Artists.* The MIT Press, 2008.
[B09]	Boudreaux, T. *Programming the iPhone User Experience.* O'Reilly Media, 2009
[AD09]	Adamson, C., Dudney, B. *iPhone SDK Development*. The Pragmatic Bookshelf, 2009.
[BW08]	Byron, L., Wattenburg, M. *Stacked Graphs - Geometry & Aesthetics.*

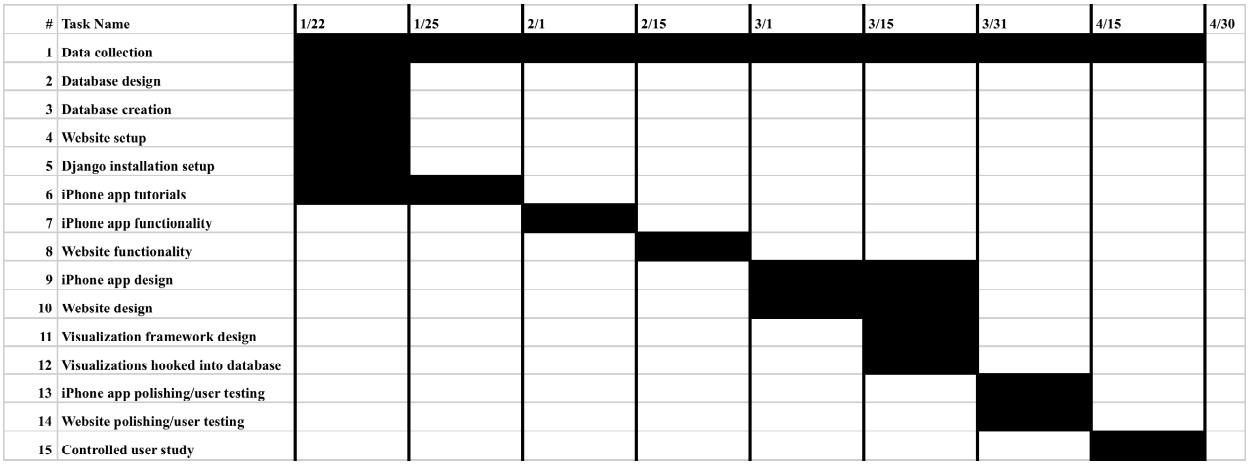

Figure 1: *Project timeline.*