SDP 02 your.flowingdata

SDP 02.1 Overview
your.flowingdata (YFD) is an application that lets people collect data about themselves and their surroundings with the popular micro-blogging service, Twitter. Twitter asks the very simple question: "What are you doing right now?" The user updates friends, family, and others in her network on what they are doing by using succinct messages, or tweets, in 140 characters or less. YFD uses this short message paradigm with a simple syntax, which in turn allows users to explore their daily habits via an online visual interface. YFD has about 3,000 users who log computer activity, music-listening, eating habits, weight, blood sugar, and many others. It serves as both a way to track short-term goals and keep a long-term data journal by intertwining data collection with everyday activity.

SDP 02.2 Approach
YFD was designed to fit in with regular Twitter habits. Millions of people use Twitter to update their network on what they are doing or what is going on around them, so the online culture of posting smaller bits of the everyday already exists. With this in mind, we can think of YFD as an extension of Twitter. Users send messages to YFD via direct messages, which are private messages, on Twitter similar they would send a regular tweet.

The point was to make data collection as easy as possible so that users could focus on analysis with the visualization tools available on the YFD site (http://your.flowingdata.com). Interaction with the data was most important in design of the site. We wanted users to be able to explore their data beyond static statistical charts. Visualization started with familiar interfaces like a calendar and tag cloud. Users can then explore their data with searchable tools for a broader view of trends, patterns, and relationships. Finally, other tools were designed specifically to highlight certain aspects of user data.

SDP 02.3 System Description
There are three main components to the YFD application: connection with Twitter, application framework, and the visualization toolset. The first component collects and parses messages from Twitter, the application framework provides user authentication and application and layout, and the visualization toolset, the focus of YFD, lets users explore their data.

Twitter
YFD has a special Twitter account that has the sole purpose to receive direct messages from users. Making use of the Twitter API, formatted messages are parsed and then stored in a MySQL database. The main challenge was creating a syntax that was flexible enough to allow different types of data collection, but at the same time was felt natural enough for users to pick up quickly.

All messages follow this format:

d yfd <action name> <value> <unit> (at) <timestamp>

Everything after <action name> is optional. An example will make this format clearer. To start simple, if a user were to track eating habits, she could use the action name of "ate." The message might looking like this:

d yfd ate

The first part, “d yfd,” tells Twitter to send a direct message to username “yfd.” The second part, “ate” is the data. The user logs when she ate. If, however, the user wanted to track not just when she ate, but what she ate, she could include value and units.

d yfd ate 2 hot dogs

In the above example, again, “d yfd” is part of Twitter’s syntax, “ate” is the <action_name>, “2” is the <value>, and “hot dogs” is the <unit>. The time of the action is automatically logged as the time of the tweet. To set the timestamp to something in the past, the user adds “at <timestamp>” to the end of her message. For example:

d yfd ate 2 hot dogs at 6:30pm
With this simple yet flexible syntax, the user can track a number of things. The user can also use hashtags (i.e. a word preceeded by a ‘#’) to categorize her data.

**Application**

Once the message is sent, it is parsed and stored. The user can then login to YFD to view her data. YFD uses the OAuth protocol for authentication, so that users can login with their Twitter username and password. The application itself was implemented in Django, the Python Web framework.

All data is kept private, unless it is made visible in the user-enabled public views. Initially, there were no public views, however, many users requested a way to share their data with others, so a way to piece together pre-selected modules for existing data was created. Some used the public views to update others on a status much in the same way people use Twitter, but from a data point of view. One user for example, updated his partner on when he was going to sleep and waking up, because he promised her he would maintain a more regular sleep schedule.

**Visualization**

YFD provides several interactive visualization tools that let users explore their data. As said earlier, users are first presented with a familiar interface in calendars and tag clouds, and then given the option to look at their data from other views. A searchable stacked area chart lets users investigate patterns and relationships over time. A treemap provides a view into aggregates, which like stacked area chart, is searchable and interactive.

Finally, there are two visualizations that display specific aspects of the data. The first shows the duration between actions. For example, the user might track her sleep patterns by sending a message when she goes to sleep and when she wakes with “goodnight” and “gmorning,” respectively. The durations visualization shows the time in between the two actions, or more specifically, the amount of time the user sleeps each night.

The second visualization explores cross-correlation between actions. Many actions that users log often correlate. As a simple example, one user keeps track of when he wakes up and what he drinks. There is a strong correlation between wake up time and when the user has his morning coffee. Similarly, the user often has a beer at the end of the day, so there is bigger time gap between wake up time and when the user drinks a beer, but the relationship is highlighted in the graphic.

Again, the main purpose of these tools is to provide users with a way to explore and interact with their data, and in turn, make conclusions about their day-to-day activities.
SDP 02.4 Accomplishments
The current version of YFD has been online and available for public use for seven months. Over 251,000 data points have been logged by 3,170 users. Many users found YFD through press from The New York Times and CNN. Others have found YFD through other Twitter users sharing with their network.

While most uses like health-related tracking were expected, there were also some unexpected uses of the service. Some made use of the Twitter API to automate data logging on YFD. Metrics included weather during a local heat wave, Web-browsing behavior and search, music-listening via the service last.fm, and computer hardware metrics such as temperature and download rate.

Twitter’s API flexibility has also allowed users to develop third-party applications to make data-logging easier on YFD. An iPhone application and Android application were developed independently along with ways to enter data with Mozilla Ubiquity and bookmarklets in Web browsers.

SDP 02.5 Future Directions
There are plans to provide a deeper visualization toolkit as well as expand the data entry syntax. YFD also provides reminders and notifications by sending direct messages back to users; however, there is still a lot of potential for “smarter” reminders. As it is now, YFD is more of a one-way application that users can send data to (with the occasional reminder). The hope is that we can expand on this to provide more feedback to users about the data they are logging and make YFD a two-way application. Ultimately we want to integrate data collection, interaction, and analysis into the everyday.