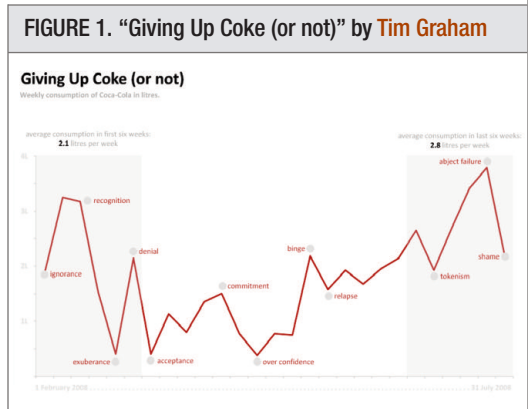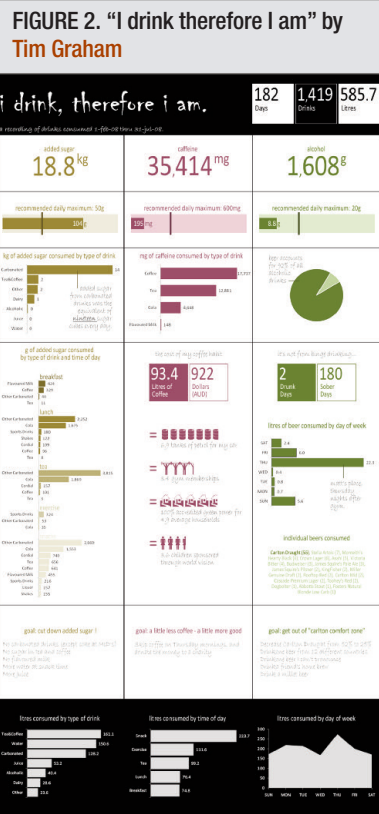# Self-Surveillance

by Nathan Yau and Jodi Schneider

Visual Representation, Search and Retrieval: Ways of Seeing

Have you ever kept a calendar, tracked what you eat or saved receipts? Simple data, like how much soda you drink, can tell a story: In fact, what you drink can tell *several* stories. Here's a more elaborate example, also by the same person:

FIGURE 2. "I drink therefore I am" by Tim Graham

FIGURE 1. "Giving Up Coke (or not)" by Tim Graham

This kind of activity is what we call self-surveillance.

What is Self-Surveillance?

What exactly do we mean by self-surveillance? It is essentially collecting data about your personal behaviors and surroundings to gain a better understanding of how you live and what goes on around you. People already upload digital pictures to *Flickr* [1], post what they are doing on *Twitter* [2] and share personal information on social networks like *Facebook* [3]. What if, in addition to these more social types of data, you were to post what you eat and drink, your blood sugar or what time you go to sleep every night? It is easy to see how such short entries about your daily habits can show behaviors over time and perhaps even a way to change the way you live for the better. We make small choices every day about what to eat, how to get to work, whom we talk to and whether to work hard or not or turn off the light when we leave a room. These tiny choices turn into habits that add up to become behaviors. Self-surveillance is a way to shine a light on these otherwise obscure and seemingly mundane aspects of our lives.

Self-Surveillance on the Web

While pen and paper work well for self-surveillance, many online applications have emerged over the past couple of years that allow individuals to collect data about themselves. Diabetics can use *SugarStats* [4] to track, monitor and share their blood sugar levels; parents can use *Trixie Tracker* [5] to collect data about their child's eating, sleeping and diaper changes; and gamers can use *Wakoopa* [6] to discover new games based on those already played. A quick survey of the available tools reveals the wide breadth of areas that self-surveillance covers as well as potential applications. Visualization is at the heart of many of these applications. It is best to keep in mind that many who log data about themselves are not data professionals. Visualization, if done right, lets users understand their data properly.

Nathan Yau is the author of the visualization blog, *FlowingData*, http://flowingdata.com, and a Ph.D. student in the statistics department at UCLA, where he works on the Personal Environmental Impact Report (peir.cens.ucla.edu).

Jodi Schneider is web librarian/assistant professor at Appalachian State University and an editor for the *Code4Lib Journal*. She follows emerging trends in visualization and the web. She can be reached at jschneider<at>pobox.com

Physical health. Self-surveillance naturally lends itself to tracking physical health and biometrics. If we think of our bodies as fine-tuned machines, then it is perhaps best we pay close attention to how we take care of them with a healthy lifestyle. Numerous studies have shown that the food we eat and the amount of time we exercise plays a major role in how we feel, and preliminary research suggests that self-monitoring increases awareness, which leads to better decisions.

In addition to the previously mentioned *SugarStats*, there are plenty of other applications to help people track metrics other than blood sugar levels. *Traineo* [7], *FitDay* [8] and *FatSecret* [9] let people keep track of their exercise routines as well as share and converse with others who are working towards common health and fitness goals. Users can track calorie consumption, weight-lifting repetitions and sets, and cardiovascular exercise time and/or distance. Weight can also be logged, as it is often the response variable that many personal trackers are most interested in. As might be expected, users can then monitor their progress via time series plots and compare their performance against others in the online community.

One common theme among *Sugar Stats*, *Traineo*, *FatSecret* and *FitDay* is that they all are manual. When someone wants to log her data, she has to go to a computer, visit the website and enter some numbers. This means that there is a gap between the action and recording. That is to say, a person will do something, write it down on a piece of paper or try to remember it and then record it on a website. But there are applications that let users log exercise data automatically. *MotionBased* [10], for example, is a product from Garmin that lets users log data with a mobile global positioning system (GPS) device that automatically uploads the data to a central server for further analysis. Nokia has a similar piece of software, *Sports Tracker* [11], which runs in the background as you go for your morning jog. Carry the phone with you as you run from point A to point B, and when you are done, view where you ran, how far and for how long from your phone or a PC. The immediate benefit is a performance report, but it should not be hard to see how *MotionBased* and *Sports Tracker* can be used to view improvement over time.

FIGURE 3. *MotionBased* (http://motionbased.com). Example from http://trail.motionbased.com/trail/invitation/accept.mb?view=dashboard&key=MTYwMg=
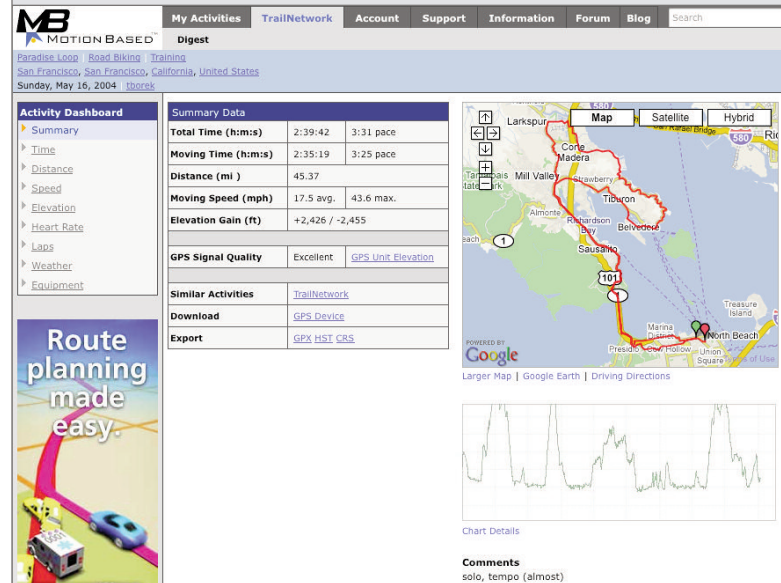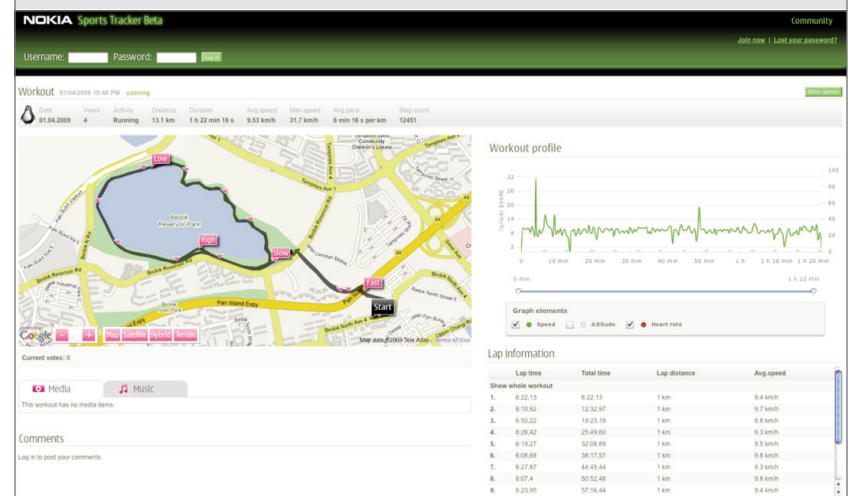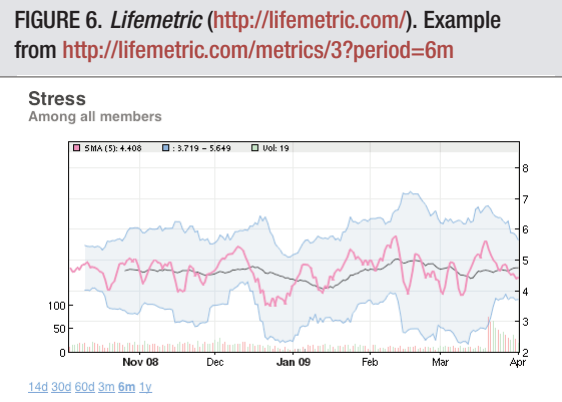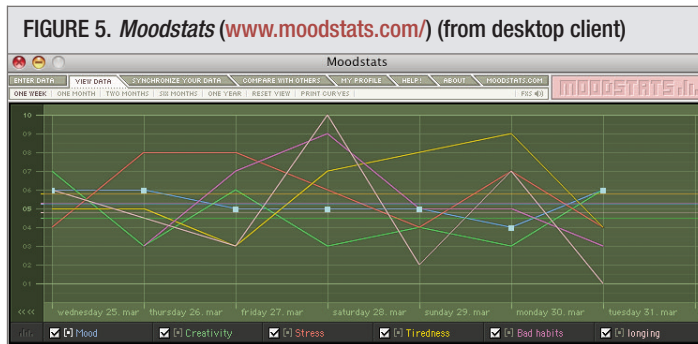

FIGURE 4. *Sports Tracker* (http://research.nokia.com/research/projects/sportstracker/). Example from http://sportstracker.nokia.com/nts/workoutdetail/index.do?id=810126

TOP OF ARTICLE   < PREVIOUS PAGE   NEXT PAGE >

Mental health. Our choices often reflect how we feel and vice versa, so it is only natural that there are personal trackers to monitor mental health, or more specifically, our emotions. Developers are challenged with the quantification of something that is innately non-numeric. What numeric value do we put on *happy*, *sad* or something in between? A few applications, while not perfect, have attempted to measure and visualize the complexity of human emotion. *Moodstats* [12] and *Lifemetric* [13] use a collection of sliders and gauges to let users put numeric values on their emotions.
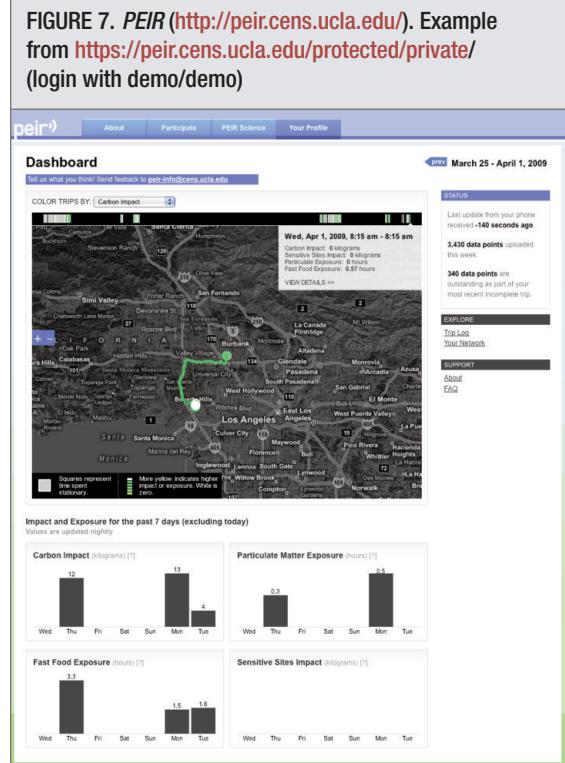
Moodstats looks and works a lot like a sound equalizer. Move the switches up and down for mood, stress levels and creativity along with some non-emotional metrics like number of emails sent, exercise and achievements. Users can also enter diary entries, which is really a form of annotation. As the user looks at past emotions, the diary can be used to make inferences about what causes certain levels of mood. *Lifemetric* is similar but also applies a social aspect to self-surveillance. Users move a scroll bar all the way left if they are sad and to the far right if they are happy and then anything in between. On any given day, users not only see how they feel, but also how the rest of the *Lifemetric* community feels.

FIGURE 5. *Moodstats* (www.moodstats.com/) (from desktop client)

FIGURE 6. *Lifemetric* (http://lifemetric.com/). Example from http://lifemetric.com/metrics/3?period=6m

Interaction with environment. Self-surveillance does not always have to be aimed inward. Sometimes it can be aimed outwards to the relationships we have with our surroundings. *The Personal Environmental Impact Report* (*PEIR*) [14] is one example. We have a give-and-take relationship with our environment that we are not always aware of. What quantity of natural resources do we consume, how much energy do we use and what effects do we have on the environment? The answers come back to the tiny choices we make every day. Do we turn the lights off when we leave a room? Do we combine driving trips to reduce travel time? These questions can go unanswered even though they can collectively have a huge impact on the environment.

PEIR allows users, based on their travel patterns, to see how they personally affect the environment in terms of carbon impact and particulate matter exposure. Users collect GPS data with their mobile phones, and *PEIR* processes the data using established micro-environment models to show users how they affect the environment and how the environment affects them.
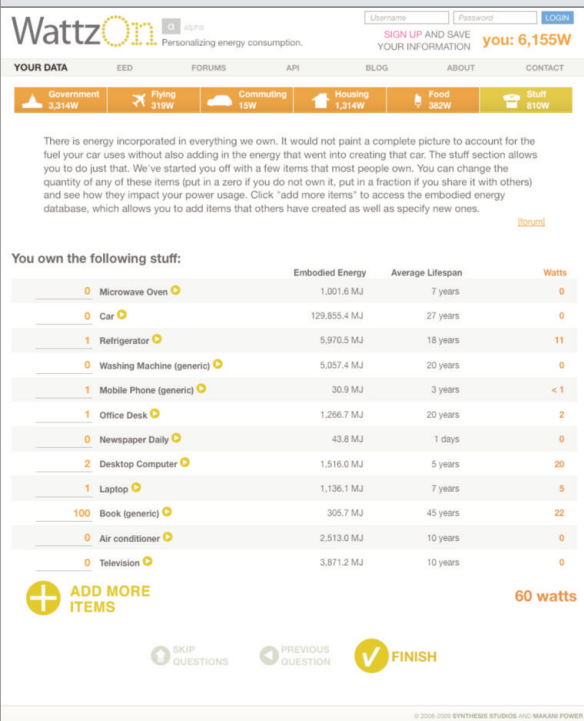
To a similar end *WattzOn* [15] uses a detailed questionnaire to estimate energy consumption. Users answer questions like how many times they fly per year on average and what kind of car they drive. Results can then be compared against others and national averages for context.

FIGURE 7. *PEIR* (http://peir.cens.ucla.edu/). Example from https://peir.cens.ucla.edu/protected/private/ (login with demo/demo)

TOP OF ARTICLE   < PREVIOUS PAGE   NEXT PAGE >

Generalized self-surveillance. Perhaps the main point so far is that if you want to track something about your life, there is probably an application that will help you do it. *Me-trics* [16], *Daytum* [17] and *Your.flowingdata* [18] are three generalized trackers that let you log data for whatever you want. Start a dataset and add to it incrementally via your PC or short messaging service (SMS) device. Some data streams are automatically logged, like number of bookmarks to the social bookmarking site *delicious.com*, while other datasets can be manual like those of the previously mentioned applications. These generalized self-surveillance tools bring several data streams into a single location, which makes it much easier to infer patterns and outliers.

The reader should notice that all the mentioned applications use visualization to present users their data. Visualization helps users, who are not necessarily data professionals, interpret their data and hopefully make educated decisions based on what they see. As a result, visualization designers have a responsibility to communicate personal data clearly in the same way that journalists have to report the news accurately. Visualizing self-surveillance data is a lot like telling a very personal story to the story's main character.

FIGURE 8. *WattzOn* (www.wattzon.com/)

It's Getting Easier

The popularity of self-surveillance is growing. As data collection and visualization advance and users become more comfortable with data, personal data collection could become common practice.

Easy data collection. Data collection technology has improved a great deal over the past couple of years – especially in the mobile arena – and our digital footprints have spread across many different streams. Our credit card bills and bank statements can be accessed online; web browsing and Google searches are stored in data logs; and grocery stores keep track of products we purchase. All these data are passive in that users do not have to actively collect it. On the flip side, there are a number of tools that let users play an active rol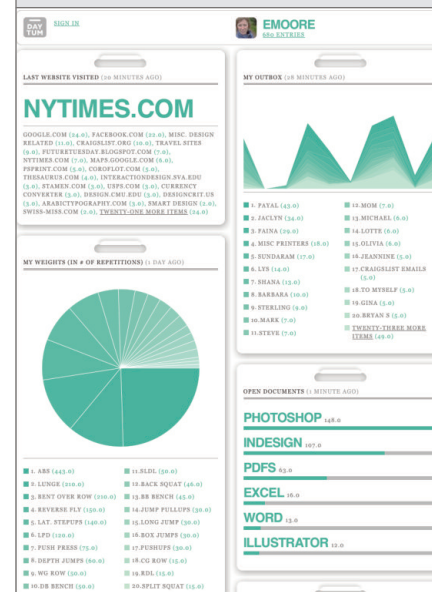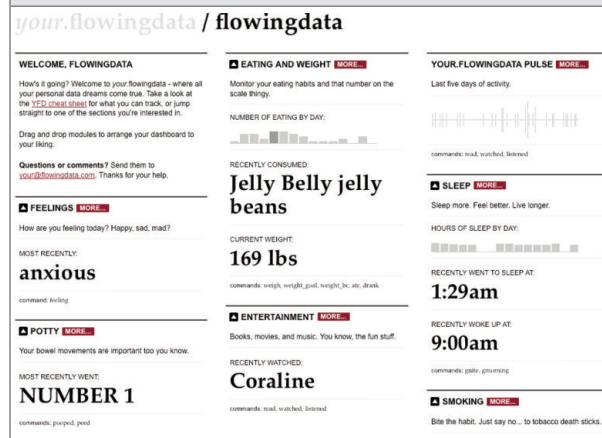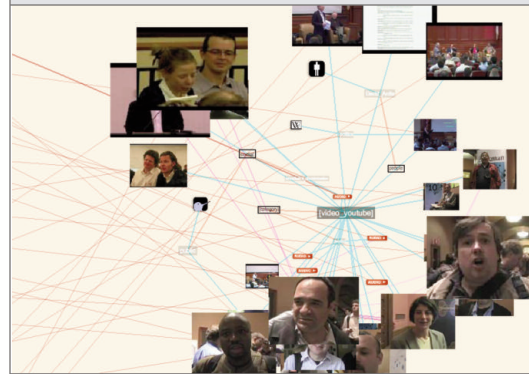e. *Carchip* [19] lets users collect driving patterns by plugging a device into a car's On Board Diagnostics II (OBD-II) port; *SleepTracker* [20] monitors waking moments with a watch; and *Fitbit* [21] measures your physical activity through the course of the day. While these tools are specifically

FIGURE 9. *Daytum* (http://daytum.com/). Example from http://daytum.com/emoore

FIGURE 10. *Your.flowingdata* (http://your.flowingdata.com/)

designed for data collection and analysis, we can also add mobile phones to the list. Millions of people own mobile phones with cameras, GPS and SMS capabilities. It is easy to see how we can repurpose devices as personal sensors.

Advances in visualization. As data collection has advanced and different streams of different data types have developed, so have the visualizations. Visualization has moved from static bar graphs and dot plots to interactive and animated visualizations. These changes, of course, have not always been for the better, but when done right, these new formats allow users to explore their data and see patterns without any formal training in analysis. Tools like *Processing* [22] and *Flash* [23] have grown more robust so that developers have more options available to them, and as a result, those who have data, have more opportunity to visualize and make inferences. Online applications like *Many Eyes* [24], from the IBM Visual Communication Lab, make visualization accessible to non-professionals.

We also see a trend in collaborations between art and science to make visualization both scientifically accurate and emotionally engaging. Works from groups like Stamen Design [25] and designers like Jonathan Harris [26] are quickly bridging the gap

FIGURE 11. *Many eyes*
(http://manyeyes.alphaworks.ibm.com/manyeyes/)

FIGURE 12. Lee Byron's last.fmstreamgraph [30]

FIGURE 13. Bestiario's social networks
(www.bestiario.org/)

between design and statistics. Such works encourage users to dig deeper and attract those who once looked at data with a skeptical eye. Works like the streamgraphs [27] [28] from Jeff Clark and Lee Byron and bestiario's social networks [29] let users explore their digital selves and have been well received with hundreds of thousands of views.

Data comfort-level. Data visualization can advance all it wants to, but not much is gained if users are not comfortable with the interface. However, we can see many examples of good design on the web. Stamen Design constantly preaches data visualization as interface and has seen much success in collaborations with sites like *Digg* [31] and *Twitter*. Most notable perhaps is *The New York Times*' willingness to try new visualization techniques. In September 2008 the *Times* graphics department produced a tree map to show market loss by major companies [32]. In August 2008 they published a weighted network graph that showed Olympic medal counts by country and year [33]. In May 2008 American spending was visualized with a Voronoi Tree Diagram [34].
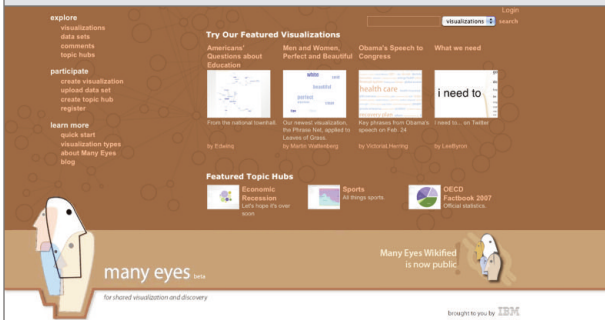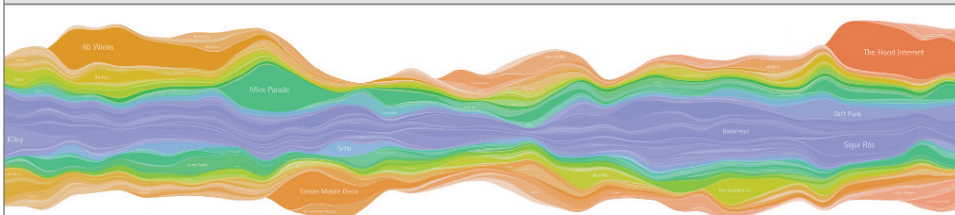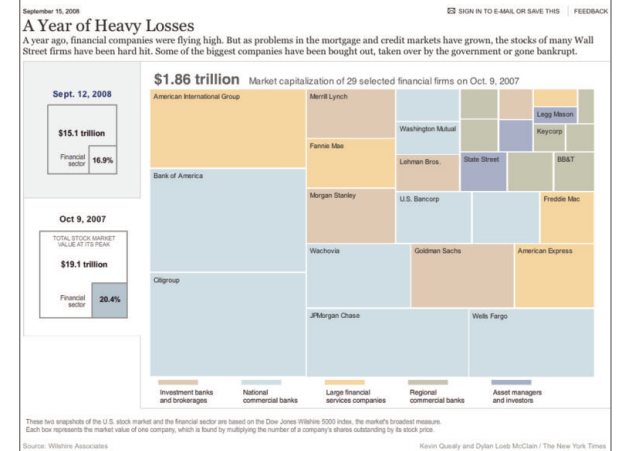
FIGURE 14. Tree map showing A Year of Heavy Losses
(www.nytimes.com/interactive/2008/09/15/business/20080916-treemap-graphic.html)

TOP OF ARTICLE   < PREVIOUS PAGE   NEXT PAGE >

FIGURE 15. Voronoi tree diagram illustrating A Map of Olympic Medals (www.nytimes.com/interactive/2008/08/04/sports/olympics/20080804_MEDALCOUNT_MAP.html)
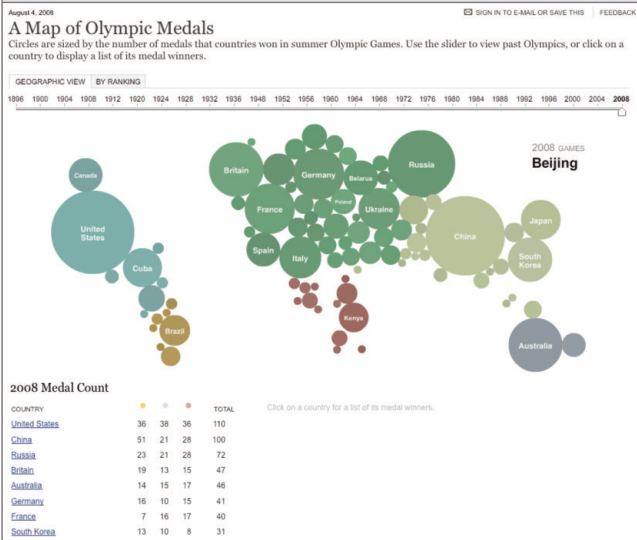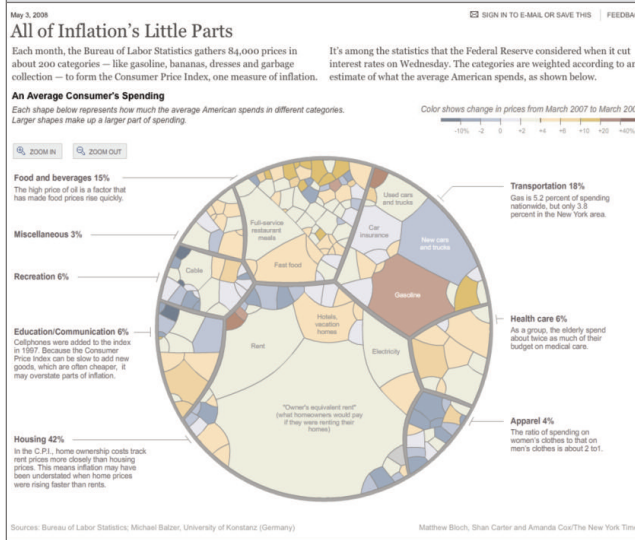

FIGURE 16. Weighted network depicting All of Inflation's Little Parts (www.nytimes.com/interactive/2008/05/03/business/20080403_SPENDING_GRAPHIC.html)

Visualization techniques once reserved for research conferences are finding themselves in national publications, and readers are saying good things about them. More importantly, though, people are exploring datasets and talking about their findings; this shows huge promise for personalized visualization. Preliminary studies show that users are more engaged when they relate to the data, and you cannot get much more personal than self-surveillance.

A Role for Information Professionals?

Information professionals can have several roles in the development of self-surveillance: advocacy, idea-generation and cross-pollination. First of all, we can advocate privacy. As individuals collect more information about themselves, often online, they rely on the safeguards and ethics of intermediaries. Information professionals can raise awareness about privacy lapses. We can also advocate for explicit data protections. For example, data portability is extremely relevant to anyone using Web 2.0 applications.

Second, information professionals can help form ideas of how self-surveillance can be used. Librarians, web developers and information designers are among those collecting and sharing their own data; this experimentation can be shared with others. We have related uses of self-surveillance in changing habits, monitoring mood, improving health and tracking our interactions with the environmental. As visualization of personal data becomes more commonplace, generalized tracking and visualization tools can be used in a variety of other ways. How? That's up to us. New ideas will come not only from our own experimentation but also from observing how data is used in other fields.

Our third role is as intermediaries. Information professionals can help cross-pollinate ideas among various communities collecting, using and visualizing data. Core concepts from library and information science, such as controlled vocabularies and taxonomies, may be relevant when summarizing data. We can share techniques from fields like statistics, e-science and cyber-infrastructure, informatics, geographic information systems and design.

Ideas to Try

Want to learn more? Try out self-surveillance for yourself. It could be as simple as writing down how much soda you drink each day this week. If you'd like to try software, but don't want to put data online, Moodstats is a good place to start: Mac and Windows users can download desktop software, create a profile and graph up to six variables. Decline synching, and check out graphs of your data after inputting numbers for three previous days. ■

Resources Mentioned in the Article

[1]   *Flickr*: www.flickr.com

[2]   *Twitter*: http://twitter.com

[3]   *Facebook*: www.facebook.com

[4]   *SugarStats*: www.sugarstats.com/

[5]   *Trixie Tracker*: http://trixietracker.com/

[6]   *Wakoopa*: http://wakoopa.com/

[7]   *Traineo*: www.traineo.com/

[8]   *FitDay*: http://fitday.com/

[9]   *FatSecret*: www.fatsecret.com/

[10] *MotionBased*: http://motionbased.com/

[11] *Sports Tracker*: http://research.nokia.com/research/projects/sportstracker/ and http://sportstracker.nokia.com/

[12] *Moodstats*: www.moodstats.com/

[13] *Lifemetric*: http://lifemetric.com/

[14] *PEIR*: http://peir.cens.ucla.edu/

[15] *WattzOn*: www.wattzon.com/

[16] *Me-trics*: http://me-trics.com/

[17] *Daytum*: http://daytum.com/

[18] *Your.flowingdata*: http://your.flowingdata.com/

[19] *CarChip*: www.davisnet.com/drive/products/carchip.asp

[20] *Sleeptracker*: www.sleeptracker.com/

[21] *Fitbit*: http://fitbit.com/

[22] *Processing*: www.processing.org/

[23] *Flash*: www.adobe.com/products/flash/

[24] *Many Eyes*: http://services.alphaworks.ibm.com/manyeyes/home

[25] *Stamen Design*: http://stamen.com/

[26] Jonathan Harris: www.number27.org/

[27] Clark, J. (2009) *TwitterStream*. Retrieved April 12, 2009, from www.neoformix.com/Projects/TwitterStreamGraphs/view.php

[28] Byron, L. (2008). *Last.fmstreamgraphs*. Retrieved April 12, 2009, from www.leebyron.com/what/lastfm/

[29] *Bestiario*: www.bestiario.org/

[30] Byron, L, & Wattenberg, M. {2008). Stacked graphs – Geometry and aesthetics. *IEEE Transactions on Visualization and Computer Graphics, 14* (6), 1245-1252. Also retrieved April 12, 2009, from www.research.ibm.com/visual/papers/leebyron_stackedgraphs_byron_wattenberg.pdf

[31] *Digg*: www.digg.com

[32] A year of heavy losses. (2008, September 15). *New York Times*. Retrieved April 12, 2009, from www.nytimes.com/interactive/2008/09/15/business/20080916-treemap-graphic.html

[33] A map of Olympic medals. (August 4, 2008).*New York Times*. Retrieved April 12, 2009, from www.nytimes.com/interactive/2008/08/04/sports/olympics/20080804_MEDAL COUNT_MAP.html

[34] All of inflation's little parts, (May 3, 2008). *New York Times*. Retrieved April 12, 2009, from www.nytimes.com/interactive/2008/05/03/business/20080403_SPENDING_GRAPHIC.html

Resources for Further Reading

Heer, J. &Agrawala, M. (2008). Design considerations for collaborative visual analytics. *Information Visualization Journal, 7*(1), 49-62.

Hesse, M. (2008, September 9). Bytes of life: For every move, mood and bodily function, there's a web site to help you keep track. *Washington Post*. Retrieved April 12, 2009, from www.washingtonpost.com/wp-dyn/content/article/2008/09/08/AR2008090802681.html

Text messaging may help children fight off obesity. (November 12, 2008). *Science Daily*. Retrieved April 12, 2009, from www.sciencedaily.com/releases/2008/11/081111142602.htm

Warren, J-B. (December 8, 2008). The new examined life. *Wall Street Journal*. Retrieved April 12, 2009, from http://online.wsj.com/article/SB122852285532784401.html