

The Buzz: Flexible Information Aggregation and Presentation for End Users

James R. Eagan
eaganj@cc.gatech.edu

John T. Stasko
stasko@cc.gatech.edu

School of Interactive Computing & GVU Center
Georgia Institute of Technology
Atlanta, GA 30332

ABSTRACT

We have presented The Buzz, a peripheral information awareness tool that allows end-users to exercise a rich degree of customization over content and presentation. We describe how users can create their own information channels to aggregate and re-present media from a wide variety of sources on the Internet.

INTRODUCTION

Every day, people send over 60 billion email messages [14], make over 1.6 million blog postings¹, and share over 32 thousand videos on YouTube [8]. Every year, more than four times the data in the entire U.S. Library of Congress is added to the World Wide Web². With such a flood of information, it can be a real challenge to keep afloat. We have created a research prototype, called The Buzz, to help people manage the deluge of information that inundates them.

The Buzz allows the user to create *information channels* to gather and present information from a variety of information sources. Additionally, users can subscribe to channels that others have already created. In this way, channels allow the user to collect information from the various publishers on the Internet and present them in a different fashion, perhaps tailored to the individual's own needs and interests. Thus, The Buzz is akin to an RSS feed reader beyond the web browser, but with the ability to create such a feed when the publisher does not make one readily available. Additionally, it uses a collage-like presentation style that emphasises images over text (although textual presentations are still available). Thus, the user can repurpose the publisher's data for use in an information awareness context.

¹As of November, 2006, Technorati.com tracks an average of 1.6 million daily posts to 60 million blogs.

²Estimated using static, publicly accessible web pages, but not private or database-driven web pages [14].

Unlike traditional feed readers, the system provides extensive support for more flexible information acquisition, aggregation, and presentation. When RSS is readily available, support exists for the user to re-purpose the data and use it in a different fashion. For example, many RSS feeds provide incomplete information, or structure the data specifically for use in a traditional feed reader, which may be at odds with the information awareness context. The Buzz provides support to extend and filter feed content, and a simple templating mechanism to allow the user to create his or her own representations of the data.

In this context, the goal of The Buzz is to help users to consolidate information from a wide variety of sources and to present each data source in a fashion appropriate for maintaining overall information awareness. As access to information increases, the demands on attention increase, making it all the more important to adequately manage our attention [5]. Peripheral awareness systems which calmly [20] present information in the periphery can help convey information without the user's explicit attention [17], helping the user to better manage the interruptions that can interfere with one's work [12].

Peripheral awareness systems have been used for a variety of other tasks, ranging from monitoring the activities of friends, family [3, 16], and coworkers [21, 11, 6], to staying apprised of information with personal relevance [19, 2]. Although some of these systems focus on expressing information from a single source, as in [3], many fall into the broader class of data aggregators ([2, 11, 19], to name but a few). These aggregators assemble data from a variety of sources for presentation.

Furthermore, information tends to be individual in nature—what is relevant to one person is often quite different from what is relevant to another. Thus, customization is vital to enabling these systems to tailor their content and presentation to support these individual differences. This tailoring can be useful for many work-oriented and personal reasons, from promoting information access to helping express an individual's identity [1].

As display costs fall, the feasibility of using displays in more ubiquitous settings increases. By moving these systems off



Figure 1. The Buzz on an extra display (right).

the desktop, exciting research opportunities arise to convey information through art [18, 7, 19], collaboration [4], and even touch [10, 13]. With such an abundance of media outlets for information, customization tools to allow the user to tailor the content and even author new information channels becomes all the more important. We use The Buzz to drive our analysis of this customization environment.

THE BUZZ

The Buzz runs on a secondary display on the desktop (see Figure 1) or on a large, public display. Each minute or so, it presents a different collage, automatically generated from some particular data source. Some existing channels include images found on the Georgia Tech College of Computing web server, news headlines from the BBC, local traffic information during rush hour, “interesting” pictures tagged by members of the College on Flickr, and current birthdays in the community. Users can select from existing channels, twiddle with basic properties of channels (such as the postal code for a weather channel), or create their own channels.

Each channel encapsulates the process of fetching the latest data from the upstream content provider, decoding that data into a standard format usable by the system, and rendering an interesting presentation to the screen. To simplify the process of decoding the content provider’s data format, a variety of data harvesters are available for various common formats, from scraping images off of human-readable web pages, to parsing machine-readable RSS, to using a content provider’s custom API. These data harvesters hide many of the details of the underlying data format from the channel creator in order to reduce the effort necessary to create a new channel.

Each channel also encapsulates the presentation aspects of drawing the data to the screen in a relatively straightforward, but flexible fashion. Various pluggable rendering modules are available for different types of data. At its simplest, the user can define regions on the screen and assign images or harvested data to those regions. The system handles the details of screen composition. In this way, even a casual com-

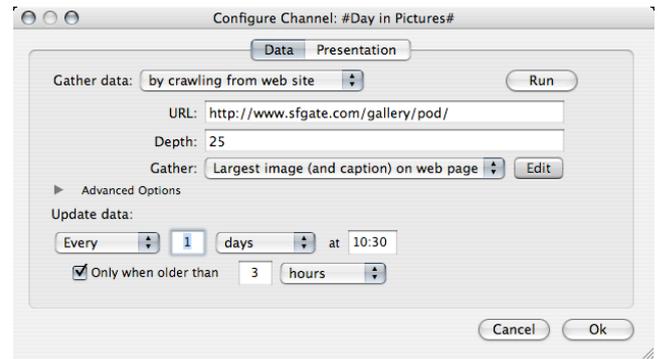


Figure 2. Data harvesting options in The Buzz.

puter user can create interesting and personally relevant information presentations.

Sharing

While traditional artforms often reflect the creative vision of just one artist, information channels often reflect the vision of several uncoordinated creators. In this fashion, information channels begin to resemble less of an artform and more of a tool. Users of customizable software, for example, rarely invent their customizations from scratch. Instead, they borrow the customizations of others and tweak them to match their own needs [15]. In this fashion, the artifacts evolve under the influence of varied creative influences.

It is important to distinguish this sharing behavior from collaboration. While collaboration necessarily involves some degree of sharing, the reverse does not. More commonly, the nature of such sharing is unidirectional. One user finds a shortcoming in the artifact of another, so she modifies it to better suit her needs. Such a shortcoming may be functional, as when a user modifies a weather channel to include the pollen forecast, or artistic, as when someone changes the layout style. Such modifications may not even reflect a shortcoming, as when one takes the Atlanta Braves baseball channel and modifies it to show Boston Red Sox scores.

All of these sharing behaviors are important to the design of such customizable media environments. Their authoring tools should, therefore, reflect this sharing behavior and make it sufficiently easy for users to exchange their artifacts with others.

EXAMPLE: THE DAY IN PICTURES

We will create an information channel from scratch in order to illustrate the process and capabilities of The Buzz. The San Francisco Chronicle publishes The Day in Pictures, a daily collection of interesting pictures found on the news wire, along with (usually) clever captions about them. Unfortunately, The Chronicle does not provide an RSS feed for these data. We will add a channel to The Buzz that fetches all of these pictures and renders them to collage-style layout.

When we create a new channel, the system presents us with a display similar to the one shown in Figure 2. In this ex-

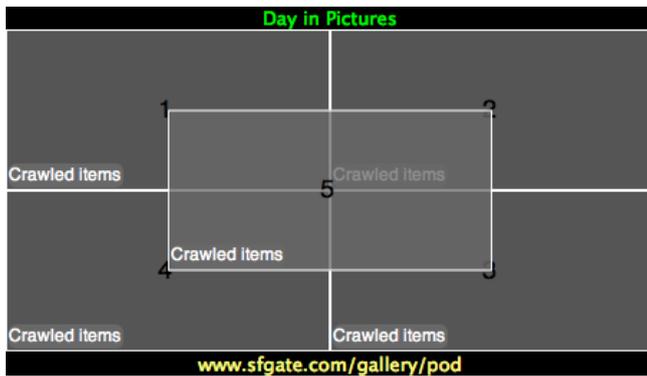


Figure 3. Day in Pictures template (left) and rendered (right).

ample, the data gathering method uses a web crawl from the SFGate website and scrapes the largest image (along with its caption) off of each crawled web page. Other possible gathering methods include using RSS or custom web APIs.

Once we have described how to gather the data, it is necessary to describe how to present the data. The default presentation uses a collage-like layout. This layout style is reasonably effective given no prior knowledge of the data to be shown. The system first divides the screen into five regions: the four cartesian quadrants and a fifth quadrant centered on top of the origin of the cartesian plane. Images are then laid out with a random jitter in each of the quadrants with the fifth image on top. In this case, the default layout works nicely for our purposes, so we will use this view.

In this way, the user can exercise creative control over the layout and placement of the various data collected and can even combine that with other images to create an appealing representation of dynamic data. Similarly, a user can quickly and easily create a simple presentation that works effectively, without needing to exert much effort. The expressibility of the system can scale with the creativity and effort of the user while still presenting an appealing presentation.

EXAMPLE: WEBCAMS

The previous example demonstrates a relatively simple example of creating a channel. Unfortunately, the various data formats in use on the Internet can complicate matters. In this example, we will create a WebCam channel. Web cams have become relatively popular on the Internet, allowing remote users a glimpse into distant and exotic worlds (or the mundane). Unfortunately, most web cams are scattered around the Internet, hosted by a wide variety of publishers, using many different representations. Some cameras publish their images to a single URL, whose contents changes periodically. Others publish to changing URLs, with a web page that dynamically updates to point to the current address. Still others encode their images in Flash or Java applications.

We will create a channel that collects various webcams from the variety of web pages on which they appear, and present them in a simple collage. The first step is to identify the we-

bcams we wish to view. We first create a new channel using the “Gather from a list of web pages” data harvester. This harvester allows the user to specify a list of URLs and various data scraping, or extraction, methods. For the webcams that publish to a fixed image URL, we can simply enter the name of the camera, its address, and use the default “identity” scraper. Other webcams are more complicated.

For webcams that publish to a web page at a fixed URL, but whose image contents update dynamically, another approach is necessary. In such a case, various heuristics may be useful. For example, being the focus of the web page, webcam content is often the largest image on the web page. As such, the user can specify that the harvester should scrape the largest image off the web page. Sometimes, however, ads vie for the distinction of being the largest image on the page, so the user can add a filter to remove those advertisements, perhaps by URL or by aspect ratio. Through these scrapers and filters, the user can tailor the extracted content.

Unfortunately, no support currently exists to handle Flash or Java. One potential approach would be to render the underlying content and scrape from an off-screen buffer [9]. Such an approach may be relatively complex and less than robust.

Finally, the user can create a layout template to display the various webcams. As before, we can create a collage-like template. This time, however, we will define an additional region that extends across the entire screen. To this region, we will bind a specific webcam that produces relatively high-resolution images of the Oregon coast. Thus, when the WebCams channel displays, it will produce a collage of five randomly-selected webcams above a dynamic beachfront view.

CONCLUSIONS

Information awareness tools offer the potential to help people manage the flood of information that inundates them. We have described a tool, called The Buzz, that enables a broad spectrum of end users to create their own customized information channel lineups, from outright borrowing of others’ channels to creating their own. Such interfaces should support a flexibility of effort that allows the user to determine how much expression he or she wishes to convey. Further-

more, authoring environments should provide explicit support for sharing and borrowing of such media artifacts.

ACKNOWLEDGEMENTS

This work has been supported in part by the U.S. National Science Foundation under grant IIS-0414667.

REFERENCES

1. J. Blom. Personalization: a taxonomy. In *CHI '00 extended abstracts on Human factors in computing systems*, pages 313–314. ACM Press, 2000.
2. J. J. Cadiz, G. Venolia, G. Jancke, and A. Gupta. Designing and deploying an information awareness interface. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*, pages 314–323. ACM Press, 2002.
3. A. Chang, B. Resner, B. Koerner, X. Wang, and H. Ishii. Lumitouch: an emotional communication device. In *CHI '01 extended abstracts on Human factors in computing systems*, pages 313–314, New York, NY, USA, 2001. ACM Press.
4. E. F. Churchill, L. Nelson, L. Denoue, J. Helfman, and P. Murphy. Sharing multimedia content with interactive public displays: a case study. In *Proceedings of the 2004 conference on Designing interactive systems*, pages 7–16. ACM Press, 2004.
5. T. H. Davenport and J. C. Beck. *The Attention Economy: Understanding the New Currency of Business*. Harvard Business School Press, Cambridge, MA, June 2001.
6. A. Fass, J. Forlizzi, and R. Pausch. MessyDesk and MessyBoard: two designs inspired by the goal of improving human memory. In *DIS '02: Proceedings of the conference on Designing interactive systems*, pages 303–311, New York, NY, USA, 2002. ACM Press.
7. J. Fogarty, J. Forlizzi, and S. E. Hudson. Aesthetic information collages: generating decorative displays that contain information. In *Proceedings of the 14th annual ACM symposium on User interface software and technology*, pages 141–150. ACM Press, 2001.
8. L. Gomes. Will all of us get our 15 minutes on a YouTube video? *The Wall Street Journal*, 248(51):B1, August 2006.
9. S. Greenberg and M. Boyle. Generating custom notification histories by tracking visual differences between web page visits. In *GI '06: Proceedings of the 2006 conference on Graphics interface*, pages 227–234, Toronto, Ont., Canada, Canada, 2006. Canadian Information Processing Society.
10. S. Greenberg and C. Fitchett. Phidgets: easy development of physical interfaces through physical widgets. In *UIST '01: Proceedings of the 14th annual ACM symposium on User interface software and technology*, pages 209–218, New York, NY, USA, 2001. ACM Press.
11. S. Greenberg and M. Rounding. The Notification Collage: posting information to public and personal displays. In *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 514–521, New York, NY, USA, April 2001. ACM Press.
12. J. M. Hudson, J. Christensen, W. A. Kellogg, and T. Erickson. "I'd be overwhelmed, but it's just one more thing to do": availability and interruption in research management. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 97–104. ACM Press, 2002.
13. H. Ishii and B. Ullmer. Tangible bits: towards seamless interfaces between people, bits and atoms. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 234–241, New York, NY, USA, 1997. ACM Press.
14. P. Lyman and H. R. Varian. How much information. Retrieved November 8, 2006 from <http://www.sims.berkeley.edu/how-much-info-2003>.
15. W. E. Mackay. Patterns of sharing customizable software. In *Proceedings of the 1990 ACM conference on Computer-supported cooperative work*, pages 209–221, New York, NY, USA, 1990. ACM Press.
16. E. D. Mynatt, J. Rowan, S. Craighill, and A. Jacobs. Digital family portraits: supporting peace of mind for extended family members. In *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 333–340, New York, NY, USA, 2001. ACM Press.
17. C. Plaue, T. Miller, and J. Stasko. Is a picture worth a thousand words?: an evaluation of information awareness displays. In *Proceedings of the 2004 conference on Graphics interface*, pages 117–126. Canadian Human-Computer Communications Society, 2004.
18. J. Redström, T. Skog, and L. Hallnäs. Informative art: using amplified artworks as information displays. In *Proceedings of DARE 2000 on Designing augmented reality environments*, pages 103–114. ACM Press, 2000.
19. J. Stasko, T. Miller, Z. Pousman, C. Plaue, and O. Ullah. Personalized peripheral information awareness through Information Art. In *Proceedings of UbiComp '04*, pages 18–35, Nottingham, U.K., September 2004.
20. M. Weiser and J. S. Brown. *Beyond calculation: the next fifty years*, chapter The coming age of calm technology, pages 75–85. Copernicus, 1997.
21. Q. A. Zhao and J. Stasko. What's Happening?: Promoting community awareness through opportunistic, peripheral interfaces. In *Proceedings of the Advanced Visual Interfaces Conference*, pages 69–74, Trento, Italy, May 2002.