2017

The Sophi HUD: A novel visual analytics tool for news media

Peter, Jessica and Szigeti, Steve and Jofre, Ana and Edall, Gordon and Diamond, Sara

The Sophi HUD: A Novel Visual Analytics Tool for News Media

Jessica Peter
Data Scientist
The Globe and Mail
Toronto, Canada
jpeter@globeandmail.com

Gordon Edall
Director, Globe Labs
The Globe and Mail
Toronto, Canada
gedall@globeandmail.com

Steve Szigeti
Visual Analytics Lab
OCAD University
Toronto, Canada
sszigeti@ocadu.ca

Sara Diamond
Visual Analytics Lab
OCAD University
Toronto, Canada
sdiamond@ocadu.ca

Ana Jofre
Visual Analytics Lab
OCAD University
Toronto, Canada
ajofre@ocadu.ca

DIS'17 Companion, *June 10-14, 2017, Edinburgh, United Kingdom*

ACM ISBN 978-1-4503-4991-8/17/06.
http://dx.doi.org/10.1145/3064857.3079151

Abstract
The Sophi Heads-up Display (HUD) is an analytics tool developed for editors at Canada's *The Globe and Mail* newspaper that overlays relevant data about articles' performances onto *The Globe and Mail* website. We describe the motivation for our research, our design process (which includes initial needs assessment through semi-structured interviews with editors, and job shadowing), and the validation of features with users. We then discuss the ongoing iterations of the Sophi HUD and how its design was informed through collaboration with *The Globe and Mail* employees.

Author Keywords
Journalism; news; visualization; heads-up display; software; design.

ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]: *user interfaces*.

Introduction
The emergence of new distribution networks coupled with increasing access to free online content has led traditional print newspapers to struggle with effectively monetizing digitally distributed journalism [3, 5, 16, 17]. However, digital platforms offer newspaper editors the ability to collect consumer data to a degree

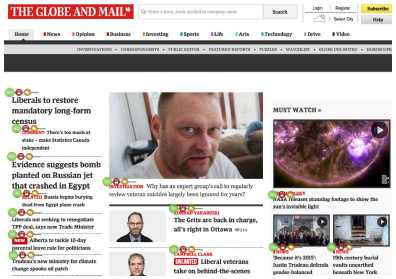

Figure 1: This mock-up includes performance indicators next to each article title.

unprecedented via print distribution. Web analytics software that presents data effectively can help users gain greater insight on readers' preferences and behaviours.

To be an effective tool for news editors, web analytics must present data that is important to editors' jobs and support the newspaper's chosen financial model(s). In this paper, we present the Sophi Heads-up Display (HUD); a visual analytics program currently under development by the authors and The Globe and Mail. Unlike many existing web analytics software programs that use a dashboard-inspired design paradigm, the Sophi HUD overlays detailed traffic data adjacent to each article link on the newspaper website. The use of data visualization software to inform the improvement of websites is well documented [6, 8, 10, 14], as is the use of a visual "overlay" [9].

Our design work, as well as our users' expectations, were influenced by existing products; including Chartbeat [4] and Parse.ly [12]. Both present a dashboard and a heads-up display system, although both were found wanting by the editorial staff.

The Sophi HUD

The HUD comprises one part of the Sophi software suite. While the Sophi suite was designed to be easy to use by staff members at *The Globe and Mail*, the HUD system has particular importance to two groups of editors; *homepage* and *section* editors.

Identifying needs and requirements

We carried out semi-structured interviews with 20 representatives of the editorial team, conducted participant observation, and undertook our design work on-site with users. Suchman's [15] call to recognize the importance of the work environment in understanding the cognitive processes when interacting with a system is central to our work. Our close interaction with editors and stakeholders provided us with the context from which the requirements emerged, allowing us to modify our initial designs with a better understanding of the challenges faced by users. We involved *The Globe and Mail* staff members throughout the development process by (i) frequently sharing iterations of our work, (ii) involving staff in addressing design challenges, and (iii) eliciting feedback through unstructured and informal meetings; an approach made possible by working directly in *The Globe and Mail* offices.

Since potential needs are diverse and users (editorial staff) number in the hundreds, the Sophi HUD must provide (i) almost instantaneous feedback regarding various performance measures, (ii) alerts regarding performance, (iii) the ability to track both short term (minutes) and longer term (hours) measures, (iv) the ability to monitor the trends of articles promoted on other sections of *The Globe and Mail*, and (v) a means to gauge overall interest in breaking topics. Due to differences in the users' familiarity with data visualizations, we opted to employ basic graphs, simple visual cues, and well-known principles of visual design.

Interface Design: Performance Indicators

The various user needs informed a series of mock-ups, including a user interface (UI) familiar to users of Chartbeat and Parse.ly (see Figure 1), and a UI which overlaid semi-transparent graphs on the article data in a split-screen context. Feedback regarding the mock-ups stressed the importance of not obscuring headlines and photographs.

Figure 2: The Sophi HUD in use on *The Globe and Mail* homepage. Note the performance indicators next to each article. The *Globe Score* is a weighted calculation that can be adjusted based on a number of variables and is intended to provide editors a holistic measure of article performance.

On the Sophi HUD, there are indicators next to each article displaying a corresponding *Globe Score*. These are colour coded to represent performance against the *Globe Score*.

A performance indicator accompanies each article link on *The Globe and Mail* homepage (see Figure 2). Each performance indicator has a line or "tail" extending to its right. The length of the tail relative to the width of the article indicates how long it has been since the article was published, capped at six hours. Colours within the tails reflect the articles' change in performance over time from right to left.

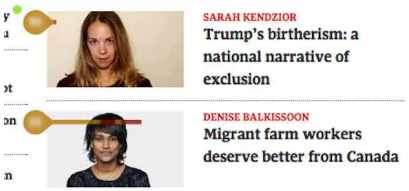

Figure 3: The colours and lengths of the article performance indicators help editors decide how to organize content. These tails have some conceptual similarity to CloudLines [11].

Users have responded positively to the "tailed indicators" given the wealth of time-based data they reflect in a limited space. For example, Figure 3 shows two adjacent articles appearing on *The Globe and Mail* website that both have Globe Scores in the "average" range. An editor may decide that, since the upper article has a very short tail, it was recently published and its middling score may be a result of the article not having found its audience yet. By contrast, the lower article has a longer tail, indicating that it has been on the website for a longer period of time. The article's tail also features a bit of green towards the middle, suggesting that the article has peaked—or, perhaps, the events it describes have passed. The editor may then consider promoting more recently published content.

Interface Design: Alerts and Charts

Editors require alerts in instances where an article is rapidly gaining or losing readers. If an article is trending in either direction, a pulsing dot is placed next to the article. The dot is coloured either green or red to indicate if the trend is positive or negative. We elected to track the rate of change in web visits over time, allowing editors to take appropriate action (for example, where an article is rapidly gaining traffic, the editor may choose to move the article to a place of greater prominence in order to capitalize on the article's popularity.) By limiting alerts, we strive to avoid a phenomenon known as "alert fatigue"; the phenomenon where a user encounters so many alarms as part of their normal workflow, that they develop a tendency to ignore all alerts [13].

When a user clicks on an article's score indicator, they are shown charts identifying related trends: a stacked area chart showing the article's performance over time, a table breaking down the number of users from different sources at different subscription statuses, a table of where the article has been promoted on *The Globe and Mail*, and a table of where the article has been promoted on social media and the number of "likes" and "shares" it has received. We opted to use tables, as they provide a good means to organize text-heavy data in relatively small spaces [18].

Additional User Interface Features
In response to editors' desire to know how content is performing in other sections, we introduced an "Elsewhere" button at the top of the page. This not only provides a means of gauging content performance throughout the website, but can also be useful for editors to see recently published content that could be promoted in the editor's designated section.

The Sophi HUD features a panel at the bottom of the page indicating the number and performance of articles above or below the user's current view. Similar to the minimap, this panel was introduced to provide a guide to how content is performing outside of what is immediately visible to the user. A number of settings are accessible by pressing the "cog" icon in the top right of the screen. This allows users to deactivate certain aspects of the screen such as removing the articles' tails, turning off flashing alerts, and turning off the panel that indicates performance measures outside of the user's current view.

Discussion and Future Work

The Sophi HUD shows article performance over time; introduces alerts on rapidly trending data; provides access to detailed article data to help the user identify the potential underlying reasons for the article's performance; and introduces a chart which provides a view of all articles published in the previous 24 hours. In order to address varying user needs, we debated the degree to which the Sophi HUD should be customizable. This poses challenges since we recognize the importance of ensuring all staff members are referring to the same statistics when judging an article's performance. But we also wanted to ensure that the HUD would meet the needs of users performing different tasks with differing goals. In describing our development of a complex visual interface, we acknowledge the key role users play from the more formal requirements gathering exercises to the numerous informal interactions between the research team and the users. The development of the Sophi HUD reinforces the critical importance of working closely with users throughout the design process [1, 2, 7, 15]. We continue to monitor the ongoing adoption of the technology in order to identify changes in work process, and potential benefits to cognition and productivity.

The Sophi HUD remains under active development. In the current news environment, it is crucial that the development of tools never ceases. We are exploring a number of features, including (i) greater customization of the widgets in order to better suit differing user tasks, (ii) increasing the complexity of visualizations for those users who are more comfortable with visual analytics, and (iii) a component that predicts whether an article should be positioned behind the paywall. As we begin formal user testing, the Sophi HUD shows enormous promise as a web analytics platform, addressing the needs of editors and other stakeholders in the world of journalism.

Acknowledgements

The authors wish to thank OCAD University's Lan-xi Dong for her research contributions, and The Globe and Mail's Jennifer Nguyen, who created the Sophi analytics system, as well as The Globe and Mail's data science team. This research is supported by the BRAIN Alliance via the Ontario Research Fund - Research Excellence (ORF-RE) program. Additional funding provided by the NSERC Strategic Parternship program, MITACS, and the OCE-VIP program.

REFERENCES

1. Graham Button. 2000. The ethnographic tradition and design. *Design Studies*, 21, 4: 319-332.
2. Bill Buxton. 2007. *Sketching User Experiences*. Morgan Kaufman. San Francisco, CA.
3. Andreu Casero-Ripollés and Jessica Izquierdo-Castillo. 2013. Between decline and a new online business model: The case of the Spanish newspaper industry. *Journal of Media Business Studies* 10, 1: 63-78.
4. Chartbeat. 2017. Retrieved February 18, 2017 from https://chartbeat.com/
5. Lesley Chiou & Catherine Tucker. 2013. Paywalls and the demand for news, In *Information Economics and Policy* 25, 2: 61-69.
6. Rik Crutzen, Johanna L. Roosjen, and Jos Poelman. 2012. Using Google Analytics as a process evaluation method for Internet-delivered interventions: an example on sexual health. *Health promotion international* 28, 1: 36-42.
7. Paul Dourish. 2006. Implications for design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '06), 541-550.
8. Wei Fang. 2007. Using Google Analytics for improving library website content and design: a case study. *Library Philosophy and Practice* (June, 2007) 1-17.
9. Tabatha A. Farney. 2011. "Click analytics: Visualizing website use data." *Information Technology and Libraries* 30, 3: 141.
10. Layla Hasan, Anne Morris, and Steve Probets. 2009. Using Google Analytics to evaluate the usability of e-commerce sites. *International Conference on Human Centered Design.*
11. Milos Krstajic, Enrico Bertini, and Daniel A. Keim. 2011. Cloudlines: Compact display of event episodes in multiple time-series. *IEEE transactions on visualization and computer graphics*, 17, 12: 2432-2439.
12. Parse.ly. 2017. Retrieved February 18, 2017 from https://www.parsely.com/
13. Shobba Phansalkar, Judy Edworthy, Elizabeth Hellier, Diane L. Seger, Angela Schedlbauer, Anthony J. Avery, and David W. Bates. 2010. A review of human factors principles for the design and implementation of medication safety alerts in clinical information systems. *Journal of the American Medical Informatics Association* 17, 5: 493-501.
14. Christopher Prom. 2011. Using web analytics to improve online access to archival resources. *The American Archivist* 74, 1: 158-184.
15. Lucy Suchman. 2007. *Human-machine reconfigurations: Plans and Situated Actions*. Cambridge University Press. 2nd ed. New York, NY.
16. Edson C. Tandoc, Jr. 2015. Why web analytics click: Factors affecting the ways journalists use audience metrics. *Journalism Studies* 16, 6: 782-799.
17. Edson C. Tandoc, Jr. 2014. Journalism is twerking? How web analytics is changing the process of gatekeeping. *New Media & Society* 16, 4: 559-575.
18. Edward Tufte. 1990. *Envisioning Information*. Graphics Press, Cheshire, CT, USA.