A Real-Time Web Observatory for Cycling Safety: A tool for supporting research and decision making of people and organizations

Roberto González-Ibáñez  Carolina Bonacic  Álvaro Fernández
Departamento de Ingeniería Informática, Universidad de Santiago de Chile, Avenida Ecuador 3659, Estación Central, Santiago, Chile
{roberto.gonzalez.i, carolina.bonacic, alvaro.fernandez.l}@usach.cl

ABSTRACT
Human-generated content on the Web has increased over the last decade. In part, this situation has been powered by the proliferation of social networking services and mobile computing. In the context of citizen safety, everyday people share information about safety-related events such as natural disasters, accidents, and theft. This information can have valuable implications for the community and authorities, however, processing and making sense of such volume of information are difficult tasks for machines and likely impossible for humans. This poster shows the design of a real-time Web observatory that uses publicly available information found in social networking services such as Twitter to automatically provide visualizations and indicators of safety-related events for the particular case of cyclists as perceived by the citizens. We demonstrate the potential applications of this design with two implementations thought to aid researchers, decision makers, and people in general.

Keywords
Information sharing, social media, web observatory, cycling safety.

INTRODUCTION
During the last decade the proliferation of social networking services on the Internet and mobile computing has eased the process to generate and share information for lay users. Despite their background, professional training, geographical location or time, people from all ages can post information online and make it available to the Internet community (Lenhart, Purcell, Smith, & Zickuhr, 2010). Everyday more and more people adopt these communication channels to share and/or acquire information (Rioux, 2000).

One of the major benefits of this evolution in online communication is the freedom of speech. Today users can exercise this right by making their opinions and thoughts available to large online communities, this by posting messages in popular online services such as Twitter and Facebook, or through other online services including blogs and news sites. With little to no filters, people can express whatever they want, whenever they want, and from wherever they are. The latter has become more common in recent years with the increasing access to mobile devices such as smartphones, which allow users to share information ubiquitously. In this sense, it is also an advantage of this form of communication the timely access to share and/or acquire information regarding emerging topics such as accidents, thefts, emergencies, crisis, and social events, among others.

On the other hand, there are several challenges regarding this massive production of content. First of all, due to the large volume of information generated worldwide, its processing is likely impossible for human beings (Schreck & Keim, 2013). For instance, getting timely access to the right information without the support of information systems can be impractical. This task can also be very challenging for machines (Samangooei et al., 2012). Discerning between relevant and non-relevant information, or evaluating the authority and reliability of content not only require computational power, but also the use of sophisticated techniques in natural language processing, image processing, information retrieval, and data mining, among others. Finally, from the lay user perspective, accessing the right information at the right time can be just luck or serendipity (Erdelez, 1997; Foster & Ford, 2002). Depending on the systems used, this can be benefited by personalization and profiling. However, for certain entities, such as government agencies, getting timely and reliable access to this information can be crucial for decision making.
This poster shows the design of a real-time Web observatory for supporting research and decision making of people and organizations in the particular domain of cycling safety. The proposed design consists of a modular and scalable solution for mining publicly available information shared by users in social networks such as Twitter and in the general Web. We demonstrate the applications of this design through two implementations (a web system and a mobile application) thought to aid researchers, decision makers and organizations in the context of cycling safety. Both implementations enable users to customize their own monitoring space (observatory), which involves the definition of information sources, filters, visualizations of geo-located safety problems as perceived by the citizens, analysis operators, indicators, reports, and alarms.

**BACKGROUND**

Social networking services are suitable means for emerging topics and rapid dissemination of information (Lerman & Ghosh, 2010). People, media, and authorities may be influenced by what happens in these large communities (Pfitzner, Garas, & Schweitzer, 2012). Today there are more than 230 million Twitter users who act passively or actively in the generation, dissemination, and evaluation of diverse content. Organizations and individuals with different motivations may be interested in monitoring this user-generated content to gain deeper access to people's opinions or even use people as sources of information. For example, government agencies are increasingly interested in using social media data for decision making and supporting the definition of public policies regarding citizen safety, which includes aspects related to disaster response, antitheft measures, terrorism, and public transportation, among others (Kavanaugh, et al., 2012). Despite this growing interest, these entities must face a number of challenges including analysis, visualization, and sensemaking of not only large amount of information, but also dynamically changing information.

**Real-Time Web Observatories**

The proliferation of popular social networking services such as Twitter and Facebook, have provided researchers and organizations with unique opportunities to investigate and mine users’ behaviors. As part of this phenomenon, online services for synthetizing and analyzing user-generated content in social media services have been developed. Examples of such services include Hootsuite¹, Twitalyzer², Brandmetric³, and Klout⁴, which are intended to help users and organizations to manage their social networks and their impact. Other services such as TweetFeel have ceased operations in part due to changes in social networking service policies. While these services offer users with a variety of tools for managing and analyzing their social networks, they are mostly general purpose. However, in certain contexts users may need specialized tools for analyzing data that is available not only on social networking sites, but also in the general Web. This includes personal webpages, blogs, and news sites, among other media that allow people to produce and share information and opinions.

Real-time Web observatories are systems capable of automatically monitoring/sensing information-related events that occur dynamically on the Web and produce meaningful synthesis of such events. A real-time Web observatory can be set up to monitor one or multiple information spaces. For example, one could create a real-time Web observatory to follow political campaigns focused on Twitter data (www.observatoriopolitico.cl) or one could create observatories for monitoring delinquency levels, transit-related issues, or population discontent, among other application scenarios.

**Cycling Safety**

Bicycles are a popular, economic, and ecologic transportation mean. Today millions use bicycles worldwide either for transportation, sports, and/or recreation. Either way, as stated by the International Transportation Form (2012), “bicycles are essential part of the urban mobility mix” (p.1).

Cycling safety is a major concern for some countries and international organizations as part of public policies related to transportation systems. Generally speaking, roads are not well-suited for cyclists (International Transportation Form, 2012), which is one of the major reasons for accidents. In spite of this situation, bicycles are usually promoted as an alternative transportation mean due to its particular positive implications in terms of health, quality of life, and economy, among other factors.

Although the formulation and implementation of public policies for transportation can be highly beneficial for cyclists, government agencies require different types of information to make decisions (Lanzendorf, M., & Busch-Geertsema, 2014; Rios, Taddia, Pardo, & Lleras, 2015). Besides official statistics that may be obtained periodically, or after long periods of time, safety, as perceived by cyclist themselves and people in general, can be a useful source for timely access to information (Wessels, Pardo & Bocarejo, 2012). In this sense the Web and particularly social networking services provide people who experience directly or indirectly safety-related events in the context of transit, with alternative means for sharing their perspective, feelings, and opinions, and disseminate this information within their networks or even to the general public. More importantly, such information is produced dynamically and sometimes in a timely fashion, which could allow others to take effective precautions to avoid accidents, assist others,

---

¹ https://hootsuite.com/
² http://twitalyzer.com/
³ http://www.brandmetric.com/
⁴ https://klout.com
and even avoid related situations such as thefts, this by knowing when and where a given event took, is taking, or will take place.

A REAL-TIME WEB OBSERVATORY FOR CYCLING SAFETY
In this section we introduce the architectural design of a real-time Web observatory for cycling safety. Then we show two particular implementations, namely, a Web-based system and a mobile app.

Architectural Design
Architecturally speaking, our observatory design is a stratified structured composed by six layers (Figure 1). The first of these layers corresponds to the Sources that will be monitored. Sources include any publicly available repository from where cycling-related data and information can be retrieved (e.g., social media sites, news sites, personal blogs, RSS feeds, etc.). Second, the Observers and Listeners layer includes processes capable of monitoring and capturing changes in the selected information sources. Third, the Filters layer corresponds to a set of procedures for selecting and cleaning cycling-related information from the feed captured by observers and listeners. Examples of filters include specific terms, regions, dates, and user accounts, among others. Fourth, the Operators layer consists of a set of functions for analyzing incoming data and produce synthesized information. Examples of operators include those for sentiment analysis, safety analysis, comparisons, seasonal analysis, geolocation, and statistics, to name a few. Fifth, the Visualization layer consist of procedures for visualizing the outputs generated by the underlying operators. Examples of visualizations include maps, heat maps, charts, timelines, and cloud tags, among others. Finally, the User Interfaces layer combines visualization components, analysis operators, filters, and resources for interacting with them. User interfaces can be implemented as Web applications, mobile applications, and windows-based forms, to name a few.

For this implementation we only used Twitter as information source. Filters were configured to capture only cycling-related events. At the user-interface level, the interface was designed to support three views or layouts (Figure 2). In the first view (Figure 2-a) the observatory displays rankings with the most used words and shared URLs. Additionally it shows a summary with the total number of tweets analyzed, positive and negative mentions, top shared pictures, and a Google map with a heat map overlaid for displaying areas where safety-related events took or are taking place. It is important to note that this user interface is automatically updated as new incoming information is captured by observers and listeners.

![Figure 2. Three views of the Web-based implementation of the Web Observatory for cycling safety.](image)

The second view (Figure 2-b) provides users with interaction means for enabling, disabling, and distributing visualization components. Additionally, users are allowed to manipulate specific filters such as query terms and analysis period. Finally, the third view (Figure 2-c) allows analyst users to perform comparisons between two periods of time to look for similarities and differences.

Mobile-based Observatory
Our second implementation consists of a mobile app for iOS. Likewise the Web-based version, this app allows users to ubiquitously interact with the cycling safety observatory. However, unlike the Web-based implementation, the mobile-app is a smart client that only implements user interface and visualizations components. Functionalities provided by each underlying layer in the architectural
design are accessed through remote Web services. A lay user or a specialized analyst can use this app to either add information that can be captured by the observatory or to obtain information from the observatory.

The app was designed for different user profiles. For instance, lay users have access to specific tools for reporting cycling-related events that will be captured by the observers and listeners of the observatory. Additionally, these users have access to visualization interfaces that show areas where cycling-related events (e.g., accidents, theft, etc.) have taken or are taking place.

Figure 3 depicts six views of the mobile-app. The home view (Figure 3-a) offers a general overview of safety-related events. The data view (Figure 3-b) provides access to samples of data used in the analyses. Visualization menu (Figure 3-c) and configuration (Figure 3-d) views allow users to parametrize the analyses and presentation of results. Finally, Figures 3-e and 3-f show two examples of visualization of geolocated data.

![Figure 3. Six views of the Mobile-App implementation of the Web Observatory for cycling safety.](image)

CONCLUSION

In this poster we introduced a novel design for a real-time Web observatory for cycling safety. We demonstrated the applicability of our design through two implementations: (1) a web-based system and (2) a mobile-app. Both systems have been tested and used in controlled environments, however the implementation of the web-based infrastructure for supporting both the Web-based system and the mobile-app were designed to be a scalable distributed system.

Both government agencies and researchers have shown interest on these tools and their potential applications. From the information science perspective, researchers are interested on this tool as a specialized resource to study information sharing behaviors of cyclists on specific contexts and locations. Those working on public policies, including researchers and government agencies, are interested on how this tool could contribute with empirical data that can be used in the definition of public policies in the transit domain.

Our future work with this Web observatory includes empirical research with specific communities of cyclists in order to study their information sharing behaviors in regards to safety-related events and how technology can be used to mine both the data and information produced by people in order to contribute to the safety of cyclists themselves and the community around them.

ACKNOWLEDGMENTS

The authors would like to thank CITTIAAPS. This article was partially funded by project PMI USA1204

REFERENCES


International Transportation Form (2012). Cycling safety: Key messages. OECD/ITF 2012


