Social Web in Disaster Archives

Michiaki Tatsubori, Hideo Watanabe
IBM Research - Tokyo
1623-14 Shimo-tsuruma
Yamato, Kanagawa, Japan
mich@acm.org

Akihiro Shibayama, Shosuke Sato,
and Fumihiko Imamura
Tohoku University
6-6-11 Aoba, Aramaki
Aoba-ku, Sendai, Japan
ashiba@saigai.str.archi.tohoku.ac.jp

ABSTRACT

Preserving social Web datasets is a crucial part of research work for disaster management based on information from social media. This paper describes the Michinoku Shinrokuden disaster archive project, mainly dedicated to archiving data from the 2011 Great East Japan Earthquake and its aftermath. Social websites should of course be part of this archive. We discuss issues in archiving social websites for the disaster management research communities and introduce our vision for Michinoku Shinrokuden.

Categories and Subject Descriptors

H.3.5 [Online Information Services]: Web-based services; H.3.7 [Digital Libraries]: Collection; K.5.1 [Hardware/Software Protection]: Copyrights

General Terms

Design, Human Factors, Legal Aspects

Keywords

social Web, disaster management, digital archives

1. INTRODUCTION

Preserving social Web datasets is crucial for research on disaster management that involves information from social media. How-ever, social media datasets grow so large that they become awk-ward to work with using the existing database management tools. The problems with social media analysis and research include capture, storage, search, sharing, visualizing, and of course the analyses themselves. Datasets also grow because they are increas-ingly being gathered by ubiquitous information-capturing mobile devices such as smartphones, which allow people to easily capture pictures and movies and store them using social networks, such as Twitter or Facebook.

This paper discusses the issues in archiving social websites and explains our vision for Michinoku Shinrokuden, an archiving project by Tohoku University. Section 2 introduces the Michi-noku Shinrokuden project. Section 3 reviews the roles of the so-cial Web during the 2011 Great East Japan Earthquake and its aftermath. Section 4 discusses the issues in archiving social web-sites for disaster
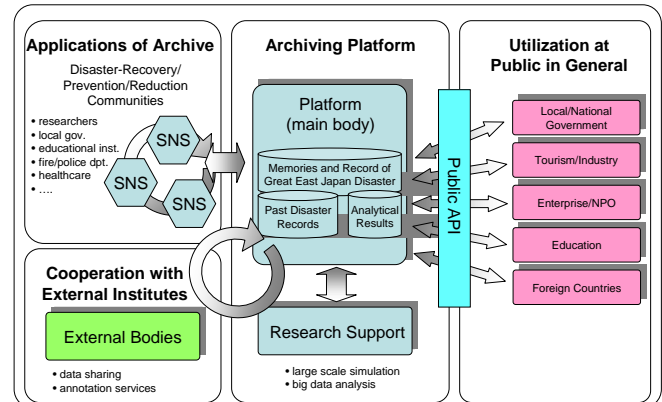
Figure 1: System architecture of Michinoku Shinrokuden (the Japan Disaster Archive at Tohoku University).

management research communities and Section 5 concludes the paper.

2. MICHINOKU SHINROKUDEN

"Michinoku Shinrokuden" is an archives project of To-hoku Uni-versity, with wide support from various groups such as content owners and ad agencies in Japan. It was launched after the 2011 Great East Japan Earthquake. Michi-noku Shinrokuden means "Record of the Michinoku Earth-quake" and refers to the region most affected by the disaster. The project is compiling archives and developing an archival system to document the Great East Japan Earthquake and its aftermath. It will collect and preserve information, rec-ollections, observations, and case studies center-ing on the March 11 events and make them available to people in other regions and to future generations. The project also aims to foster research on disaster prevention and mitigation and to pro-mote measures to prevent disasters and to reduce their impact throughout Japan.

While the archiving targets for Michinoku Shinrokuden in-clude a wide variety of data sources such as TV broadcasts, newspapers, sensor data, and interviews, this paper focus on social websites. The system is also intended to be used with Social Network Sys-tems (SNSs) for disaster manage-ment. Figure 1 illustrates the abstract system architecture of Michinoku Shinrokuden.

Here are 8 of the guiding principles (out of 11) for the disaster archive project:

- Collect a wide range of information such as anecdo-

tal ac-counts, records, case histories, and knowledge without limit-ing the possibilities for their future uses.

- Archive earlier disasters in the Tohoku region and pre-pare to record future events, in addition to archiving the records of the Great East Japan Earthquake and its related events.

- Disseminate the information on disaster responses as needed in forms that support disaster prevention and responses.

- Allow the archive system to grow and evolve in re-sponse to the interactions with its users.

- Publish the current status of ongoing rehabilitation and recon-struction work after the disasters (with a chronology) with a proper coverage (of the geography) from the coast to inland parts of the Tohoku region, in a quasi-real-time manner.

- Cooperate with other institutions with earthquake dis-aster archives used for social services.

- Collaborate with disaster prevention and response sys-tems at universities, research institutions, government agencies, and private companies to support their dis-aster relief preparations and countermeasures.

- Provide educational information on disaster preven-tion and mitigation in various environments (such as e-Learning).

3. 3/11 DISASTERS AND SOCIAL MEDIA

The 2011 Tohoku earthquake, also known as the 2011 Great East Japan Earthquake, caused many types of prob-lems. The degree and extent of the damage caused by the earthquake and the related tsunami were enormous, with most of the damage being caused by the tsunami.

Researchers reported that Web-based social media such as Twitter and Facebook were heavily used after the 2011 Great East Japan Earthquake [5, 8, 2]. While this was partially due to the unavailability of more informa-tive websites and confusion related to the nuclear power station disaster, we believe that it was mainly because such media are becoming very popular in people's lives and because they have many characteristics suitable for such situations.

When the March 11 earthquake struck, Facebook users globally (in dozens of languages) posted updates, videos, photos, news links, donation requests, and millions of prayers and words of sympathy, beginning almost immediately after the earthquake. Facebook was reported to have approx-imately 4.5 million status updates from 3.8 million Face-book users around the world on March 11 that mentioned "tsunami", "Japan", or "earthquake." [6]

Private individuals and groups used Twitter for their own communications. This included 'street journalism', news, alerts, or opinions from many people who are not profes-sional journalists. This form of public journalism has two forms. First, the participat-ing 'journalists' sent reports, photos, videos, or information to commercial news sites. Second, citizen journalists developed their own news con-tent and posted their unedited products on their personal websites. These sites may be managed by their own users, or individuals may be posting information to websites hosted by others [1].

4. ISSUES IN SOCIAL WEB ARCHIVING

On of the biggest issues in archiving social Web is the copyrights for the social content. Redistributing social con-tent from an ar-chive without the permission of the content's creators violates the copyright laws of most countries. We need methods to limit the use of archived content to research or other forms of fair use.

Another problems involves multimedia content, especially photos and movies. Important information is included in the social Web [3]. However, without textual metadata for indexing such content, it is difficult to search in photos and movies. We are test-ing techniques for semantic modeling and classification of images and video content [7] with crowd sourcing for collabora-tive tagging through the "wisdom of crowds" [4].

5. CONCLUDING REMARKS

Preserving social Web datasets is a crucial part of research work for disaster management with information from social media. In this paper, we introduced the Michinoku Shin-rokuden disaster archives project and discussed issues in archiving the social Web.

6. REFERENCES

[1] C. J. J. The great eastern japan earthquake - assess-ing disaster response and lessons for the united states, 2011. http://report.heritage.org/sr0094.

[2] B. R. Lindsay. Social Media and Disasters : Current Uses , Future Options, and Policy Considerations. CRS Report for Congress, R41987, 2011.

[3] S. B. Liu, L. Palen, J. Sutton, A. L. Hughes, and S. Vieweg. In Search of the Bigger Picture : The Emergent Role of On-Line Photo Sharing in Times of Disaster. In Proceedings of the 5th International ISCRAM Conference, Washington, DC, USA, May 2008, number May, pages 140–149, 2008.

[4] S. Sen, S. K. Lam, A. M. Rashid, D. Cosley, D. Frankowski, J. Osterhouse, F. M. Harper, and J. Riedl. tagging, communities, vocabulary, evolution. In Proceedings of the 2006 ACM Conference on Computer Supported Cooperative Work, CSCW 2006, Banff, Alberta, Canada, November 4-8, 2006, pages 181–190. ACM, 2006.

[5] F. N. Shigyo. The Great East Japan Earthquake: How Net Users Utilized Social Media? The NHK Monthly Report on Broadcast Research, 61(8):2–13, Aug. 2011.

[6] R. Tsuchiyama. The great eastern japan earthquake and social networking in japan. a blog on April 20, 2011, http://www.forbes.com/sites/raytsuchiyama/2011/04/20/the-great-eastern-japan-earthquake-and-social-networking-in-japan/.

[7] L. Xie, A. Natsev, J. R. Kender, M. Hill, and J. R. Smith. Visual memes in social media. In Proceedings of the 19th ACM international conference on Multimedia - MM '11, page 53, New York, New York, USA, Nov. 2011. ACM Press.

[8] Y. N. Yoshitsugu. Roles of social media at the time of major disasters observed in The Great East Japan Earthquake: twitter as an example. The NHK Monthly Report on Broadcast Research, 61(7):16–23, 2011.