Save the tweets so you can understand the birds

[Extended Abstract]

Claudia Serbanuta
University of Illinois
501 E Daniel Street
Champaign, Illinois
cserban2@illinois.edu

Tiffany Chao
University of Illinois
501 E Daniel Street
Champaign, Illinois
tchao@illinois.edu

Aiko Takazawa
University of Illinois
501 E Daniel Street
Champaign, Illinois
aikot@illinois.edu

ABSTRACT

The emergence of Web 2.0 technologies in the digital age offers increased opportunities for worldwide information dissemination and self-expression. However, the capture, archive, and future use of this user generated content as a scholarly resource has been seldom addressed in the either the library or archive context. This poster presents a case study regarding the virtual exchanges and events that followed the April 2009 Parliamentary elections in Moldova. The civil unrest in the ex-Soviet republic was first presented in the international media as a "Twitter revolution" and the buzz 'web 2.0' label stayed with the events as they garnered the increased attention of a global public (see references).

The events, the information flow and the emotional outpouring surrounding the Moldova election were primarily manifested through Twitter. The community-driven use of this online platform to convey instantaneous coverage illustrates the increasing use of digital spaces as a public extension for social interaction. Introduced in 2006, Twitter has become one of the most popular internet applications for personally conveying brief textual messages to a group of potentially interested individuals (microblog); preliminary research has been conducted regarding the intentions for using this type of microblog by individuals and the communities formed (Java, 2007). The aggregation of these messages (tweets) based on a common identifier (hashtag), which links with an issue or event, presents an invaluable record of social memory with immense contextual and research value.

Tweets containing the hashtag created for Moldova events (#pman) were archived by the authors. Using this raw data, we will explore the value of such information from an interdisciplinary, scholarly perspective. By analyzing the #pman archive, this poster will address the following questions: What does the content in the archived messages tell us about the events from Moldova and the people involved? Who might use such an archive of user-contributed content in the future? How is preservation of this type of digital content made possible?

Through the investigation of the archived Twitter content, we will explore what these messages tell the public about the events in Moldova. We will look at the depth of data contained in tweets and in the meta-information available through Twitter. The content will be examined for information on how communication and information sharing transpired. Special attention will be given to: the language used (Romanian, English, or Russian) in the same online space to describe the events; the structure of the messages (how the statements were constructed and how they circulated) and hyperlinks (the outside information sources that were attached to the posted messages, i.e. URLs to multimedia sources). Patterns, observations and further areas of inquiry will be presented in the attempt to establish the value of preserving such a unique data set.

Based on its social and contextual richness, this data set would be of interest to scholars from a variety of fields. During the election and subsequent protests, this content became part of not only the local but also the global community memory. It also serves as an essential resource which documents the entire scope of events for future users to confront and interpret. Our analysis of the Moldova election response demonstrates a potential application for the development of a civic society by recognizing community memory and identity, and its lasting impact on social participation and democracy. From this perspective, the individual messages and images shared have sustainable value in the aftermath of such intense social and historical events. Inherited values of the community as well as memories of the event remain alive and continue to emerge at different levels of society as well as in individual thought, patterns which could be potentially traced through the preserved record of the #pman archive.

With the limitation of 140 characters, tweets and other types of user-generated content in Web 2.0 environments have generally been considered ephemeral in nature where long term preservation of such information has not been observed. Our case study illuminates an opportunity to learn more about the preservation needs for such unique content by examining the process and the types of information that have been captured (i.e. time stamp for when tweet was posted, native language of the tweet) while also considering the additional details and components that would make this data set more robust for potential reuse by another researcher. The archiving initiatives of electronic mail that have been undertaken by government, institutional organizations, and repositories along with the emergence of online archival services that specialize in Twitter-produced content...
(i.e. www.twapperkeeper.com) will serve as a comparative framework for determining the preservation criteria of our dataset (Lukesh, 1999).

As part of the microblogging and Web 2.0 phenomena, Twitter presents a crucial reflection of modern social culture in the digital realm. Our exploratory case study of the Moldova "revolution" shows the importance of this online venue in its capability to document informal exchanges from a worldwide network of participants on a variety of social, political and cultural issues and events as they occur and unfold. The content generated through Twitter helps to strengthen the knowledge of specific historical and social contexts if attention is directed to its preservation. Investigation of the #pman archive will reveal the depth of community participation and the impact of such a resource for long-term study while introducing a new primary source for preservation consideration.

Following the Moldova elections in April, another "Twitter Revolution" complemented the Iran election protests in June 2009. The international media and public reactions were even more prominent and pervasive- this time, tweets were more efficient in carrying these messages around the world as seen in the increased number of internet users archiving tweets with #iran and #iranelections. Through the construction of these archives and the provision of public access to them, the inherent value of this content is not only reflective of those thoughts and emotions expressed by the people behind the tweets but also of those individuals who sought to protect the ideas they embody.

Categories and Subject Descriptors
H.4.3 [Information Systems Applications]: Communications Applications; H.3.5 [Information Storage and Retrieval]: On-line Information Services—Web-based services

General Terms
Documentation, Reliability

Keywords
Twitter, digital archives, social networking

1. REFERENCES
http://www.rferl.org/content/Moldovas_Twitter_Revolution/1605005.html Accessed on 8 April 2009.
[5] S. Lukesh. E-mail and potential loss to future archives and scholarship or the dog that didn’t bark. FirstMonday, 4(9), 1999.