Stuff Happens Continuously: Exploring Web Contents With Temporal Information

Omar Alonso
Microsoft Corp.
1065 La Avenida, Mountain View, CA 94043
omar.alonso@microsoft.com

ABSTRACT

In the last few years there has been an increased interest from researchers and practitioners in exploring time as a dimension that can benefit several information retrieval tasks. There is exciting work in analyzing and exploiting temporal information embedded in documents as relevance cues for the presentation, organization, and the exploration of search results in a temporal context.

Most of the current approaches focus on leveraging the temporal information available in document sources like web pages or news articles. However, the Web keeps evolving beyond simple web pages and new information sources and services are adopted very rapidly. For example, the incredible amount of content that is generated by users in social networks offers another aspect to examine how people produce and consume content over time.

We review the current activities centered on identifying and extracting time information from document collections and the applications to the information seeking process. We outline the potential of new sources for studying temporal information by presenting new problems. Finally, we discuss a number of scenarios where a temporal perspective can provide insights when exploring Web contents.

Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous; H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval

Keywords

Temporal information, timelines, information retrieval, social.

1. CONTENTS

Enthusiasm on temporal information is on a continuous rise in the community. This new interest motivated the organization of two workshops at top information retrieval conferences to foster expertise sharing and promote early research [1, 3]. This year there was a comprehensive tutorial on temporal dynamics [2] and a TREC is organizing a new task on temporal summarization [4].

Timelines are becoming more mainstream in many applications as a way of displaying a list of events in chronological order. The quality of tools for recognizing time is improving and users value time for arranging information. The time is right to engage in this exciting field of information management.

2. REFERENCES

[1] Fernando Diaz, Susan T. Dumais, Kira Radinsky, Maarten de Rijke, Milad Shokouhi (Eds.) "Workshop on Time-aware Information Access (#TAIA2012)", *ACM SIGIR* 2012.

[2] Kira Radinsky, Fernando Diaz, Susan T. Dumais, Milad Shokouhi, Anlei Dong, Yi Chang. "Temporal Web Dynamics and its Application to Information Retrieval", *WSDM*, pp. 781-782, 2013.

[3] Temporal Web Analytics (http://temporalweb.net/)

[4] TREC Temporal Summarization Task (http://www.trec-ts.org/home)

Short Bio

Omar Alonso is a technical lead in Microsoft. He holds a PhD in computer science from the University of California at Davis.