

Connecting Documents to Ideas

Beth Trushkowsky
Computer Science Division
University of California at
Berkeley
trush@berkeley.edu

Rob Ennals
Intel Research
2150 Shattuck Avenue
Penthouse Suite
Berkeley, CA 94704, USA
robert.ennals@intel.com

ABSTRACT

Much of the information contained on the web consists of opinions, arguments, blogs, articles, and other such unstructured natural language text. Each such document contains one or more statements or ideas that might be of interest to readers.

We propose a new approach to finding, organizing, and sharing ideas contained in web pages. In this approach, a user marks points of interest

1. INTRODUCTION

IDEA: Much interesting data on the web is unstructured text

IDEA: These documents consist of ideas

IDEA: Think Link connects documents to the ideas they express

IDEA: This paper is itself marked up in think link - with URL

IDEA: If the ideas in a document are known then this allows interesting queries

2. WHY PEOPLE ANNOTATE

IDEA: People already annotate documents, in the form of blogs, comments etc

IDEA: People already organize data, using text files, Google Notepad, OneNote, etc

IDEA: People already share information, by mailing links, posting stories on facebook, sending messages on twitter, etc

IDEA: People already want to see what each other are interested in - using Facebook, Twitter, Friendfeed, etc

STUDY RESULT?: Our users used these tools

STUDY RESULT?: Users want to put forward their opinions

STUDY RESULT?: People want to be seen as a source of useful information

STUDY RESULT?: People want to organize the information they find

STUDY RESULT?: Our users said they would use our tool

STUDY RESULT?: The best way to find answers is social - by learning from friends who understand the topic

3. ORGANIZING KNOWLEDGE

IDEA: Like Wikipedia without the articles

IDEA: Organize ideas with Points, snippets, and topics

IDEA: A point can support or oppose another point

IDEA: A point can be exactly the same or exactly opposite

IDEA: Important to have wiki features to keep things in check

STUDY RESULT?: People found it useful to organize ideas

STUDY RESULT?: People found the topic approach worked well

STUDY RESULT?: People were able to easily find and reuse existing points and topics

4. USING THIS KNOWLEDGE

IDEA: Show me the points that other people disagree with

IDEA: Show me the points that other people found interesting

IDEA: Show me the points that my friends found interesting

IDEA: Show me the articles that make interesting points I haven't heard before

IDEA: Show me when new snippets appear on points I'm interested in

IDEA: Show me when new points appear on topics I'm interested in

STUDY RESULT?: People liked to be able to monitor interesting new ideas on topics they were interested in

STUDY RESULT?: People liked the other features too

5. CREATING KNOWLEDGE

6. SHARING KNOWLEDGE

Why do users g

7. USING KNOWLEDGE

8. QUERYING DATA

9. RELATED WORK

9.1 Web Annotation Systems

9.2 Semantic Linking

9.3 User Contributed Links

Memex allowed users to describe their own paths between pages, rather than requiring the page authors to modify them.

Everything2 creates soft-links between documents based on the browsing patterns of users.

ENQUIRE contains bi-directional links that show who links to a document. TrackBack does the same.

9.4 Automated Reasoning

9.5 Argumentation Graphs

10. EVALUATION

11. CONCLUSIONS