

Blogging with CONTEXT:

A context-aware information retrieval system for bloggers

Anatoliy Gruzd – Dalhousie University – gruzd@dal.ca

Justin Wong – University of Toronto – jbh.wong@utoronto.ca



Objectives

The poster describes the development process and evaluation of a context-aware information retrieval system for blog authors, called CONTEXT.

As a blogger writes a post, the system automatically retrieves and displays links to web resources (web sites, news stories, video, etc) that are relevant to the blogger's entry. CONTEXT continuously monitors the blog post for new search keywords to find and display the most relevant resources to the blogger.

Goals

- Provide real time information on related discussions in the blogosphere
- Facilitate greater participation in 'global' conversations
- Improve blog content quality through exposure to others' ideas and opinions

User Interface

The user interface of CONTEXT adds to the 'Write New Post' or 'Edit Post' page within Wordpress, introducing another pane below the main writing area.

Overall, the user sees a two part interface when writing or editing a post: the upper half of the screen for writing or editing, and an additional pane on the lower half of the screen, which contains the controls and display for CONTEXT.

The lower pane is separated into a tabbed interface displaying the search results of different web resources: other blogs, video results, and news results.

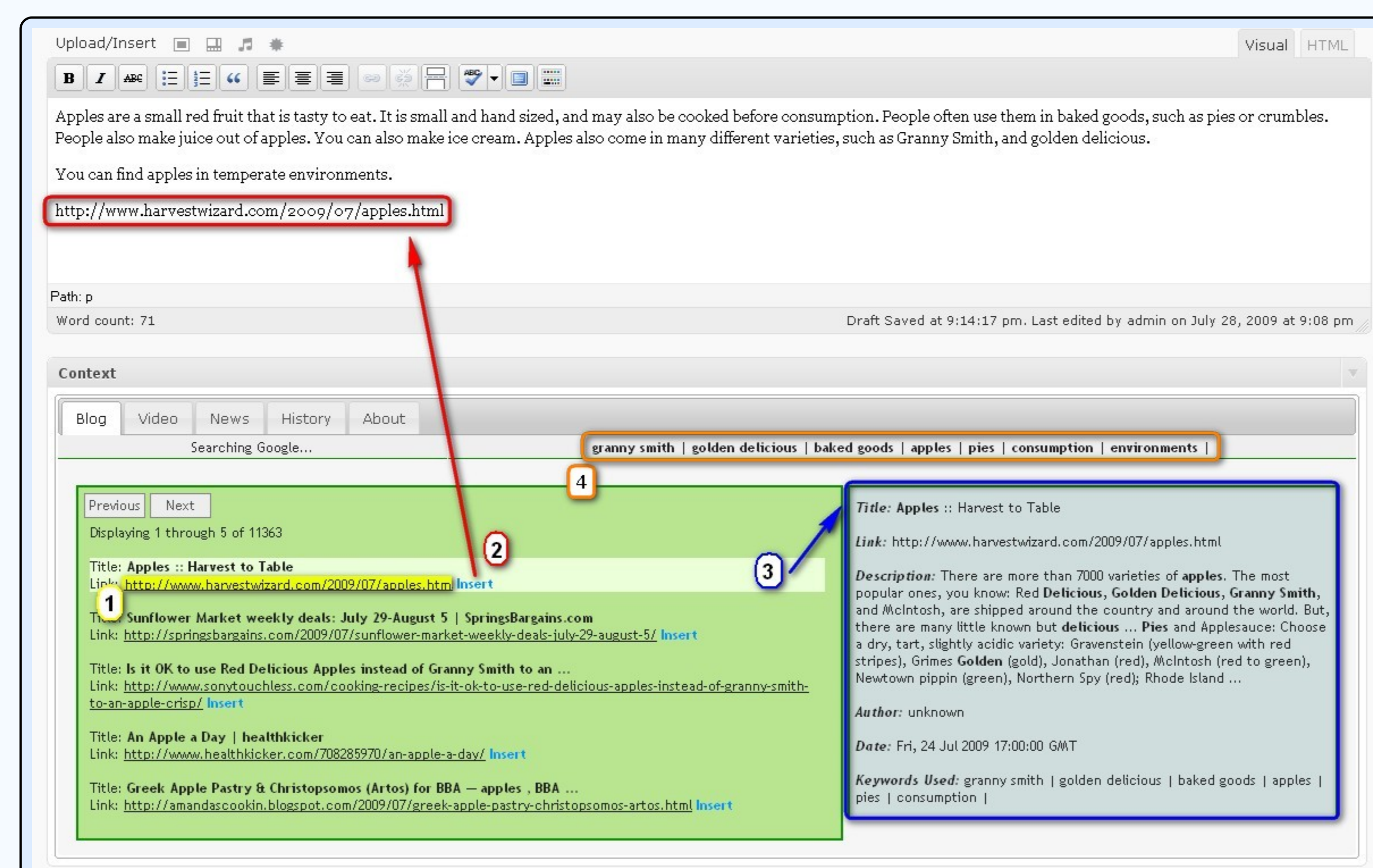


Figure 1 - Main Interface

Search results are continuously updated as the user's blog entry changes. There are several points of interest in the interface (See Figure 1):

1. URL of a search result. The user may click to view the resource.
2. Users may click "Insert" to add the URL directly to blog text
3. Moving the mouse over a title displays detailed information about the result.
4. Display of keywords extracted from current text.

Architecture

At its core, CONTEXT is a plug-in for the popular blogging platform Wordpress that relies on third party web services for searching and text analysis.

CONTEXT actively monitors the main content box of a posting or editing page. While users type, the text of the content box is analyzed for relevant keywords using the Yahoo! Term Extraction Service.

Keywords are queried in an external search engine and the results along with keywords, are displayed on screen.

Figure 4 (See right) describes the process of functionality within CONTEXT.

Currently, CONTEXT relies on Google's Application Programmable Interface (API) to retrieve relevant results from three popular types of information resources in the blogosphere: other blogs, news stories and videos.

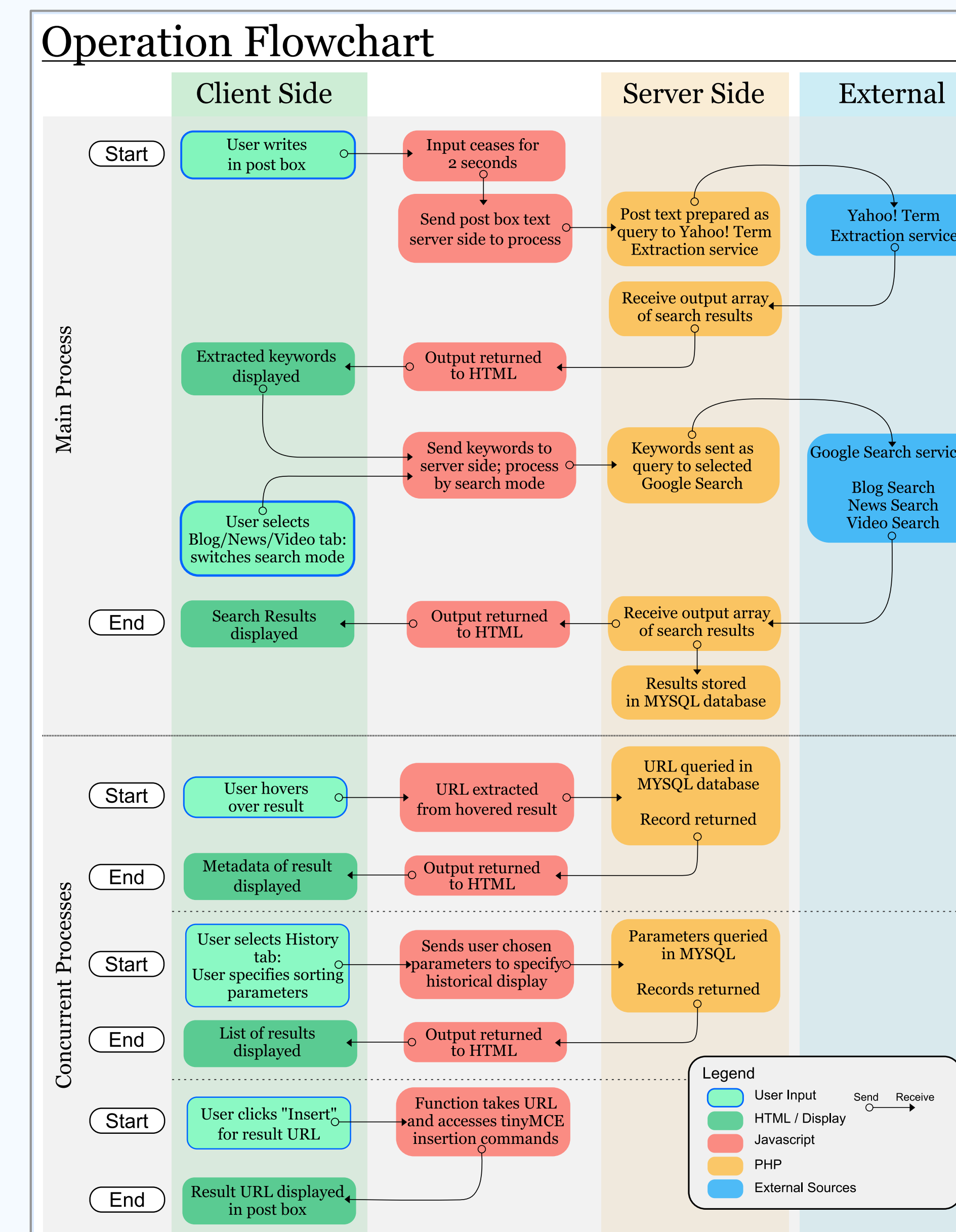


Figure 4 - Process Flowchart

Users may change the default settings of CONTEXT to specify the type of search results they wish to review first. These settings are accessed through the main sidebar, as with other Wordpress plugins. (See Figure 2)

Users also have the option of viewing past search results via the History tab. These results can be sorted in several ways; for instance, the date which the resource was first published on the web, type of search media (regular, news or video) (See figure 3)

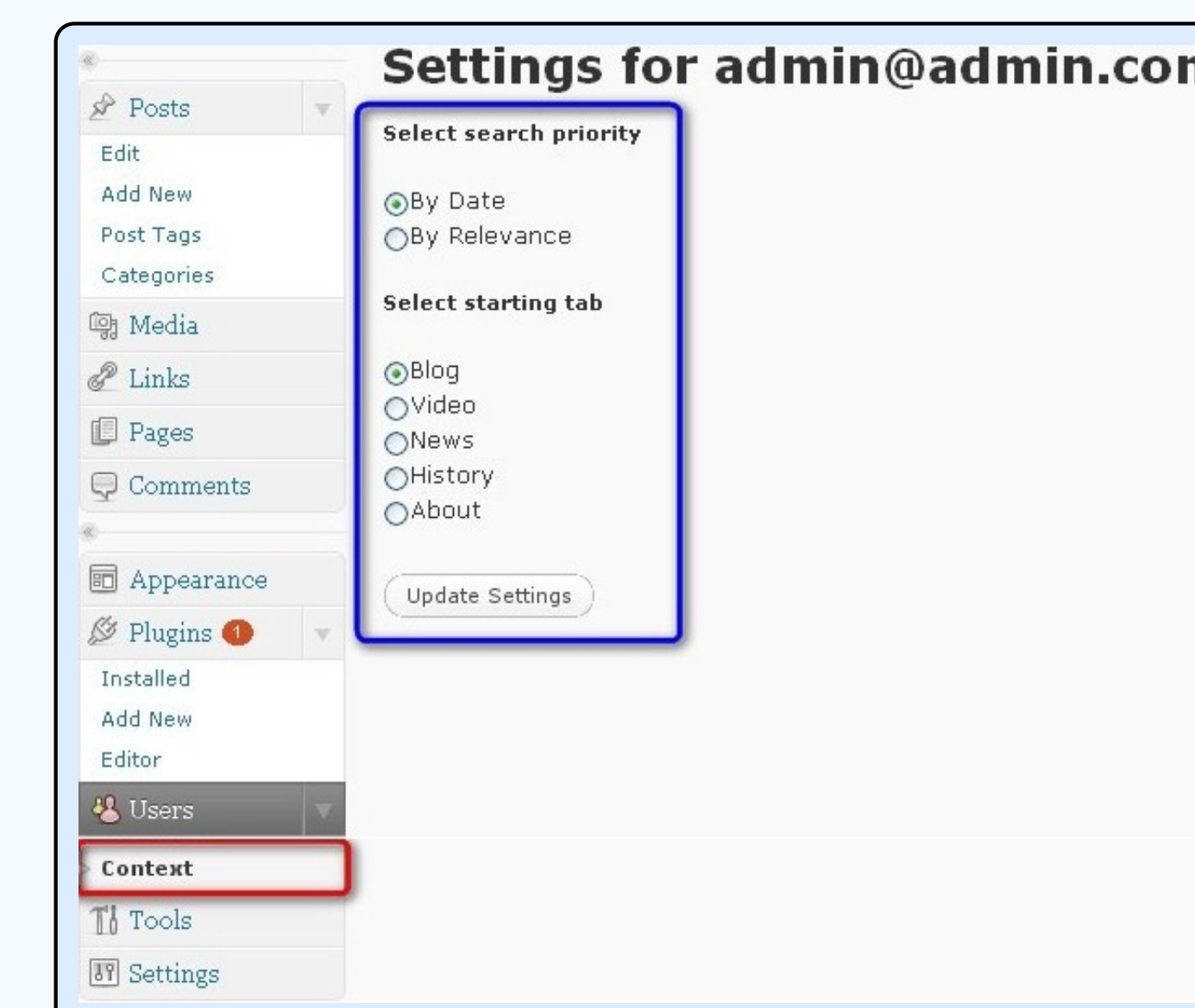


Figure 2 - Preference Settings

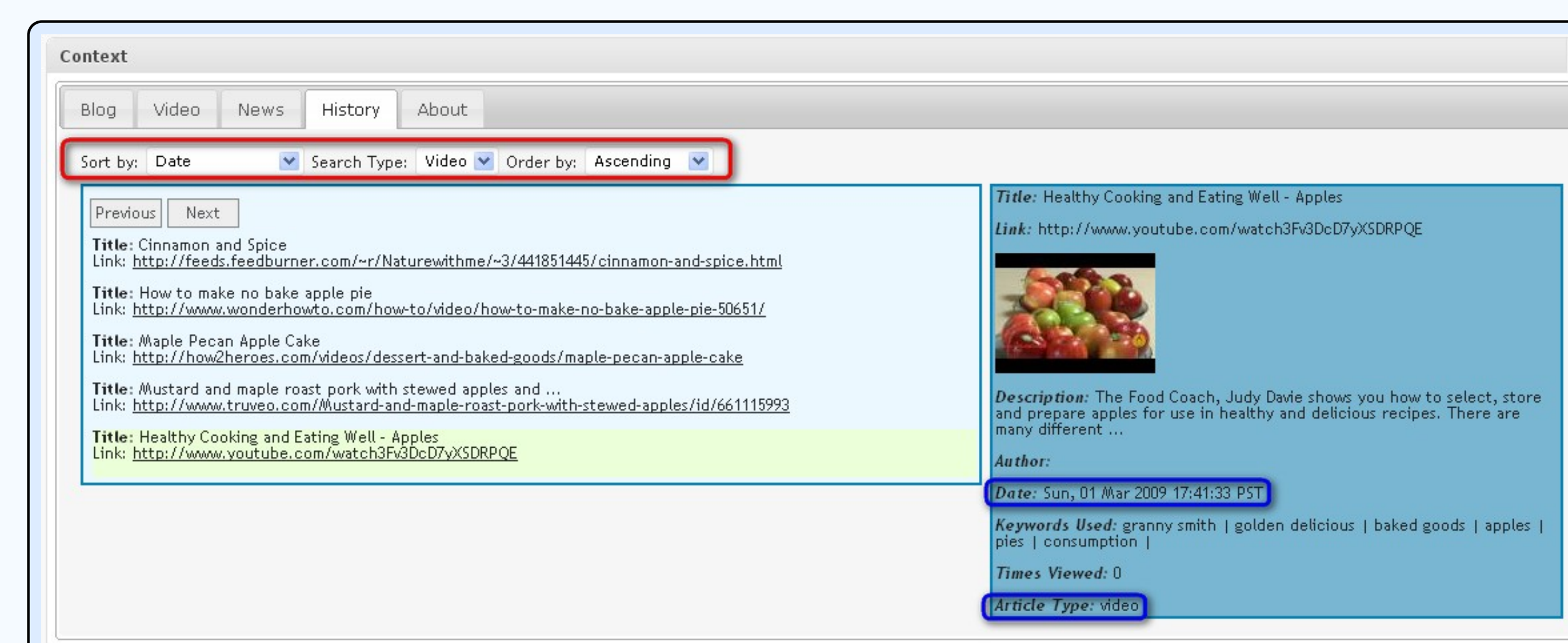


Figure 3 - History Tab

Proposed Study Design

We intend to invite up to 50 potential users to participate in preliminary usability study. Participants are required to have maintained an active weblog for at least 3 months prior to the study.



Study participants will be randomly divided into two groups and assigned to write about the same two topics. Each group will perform the same Writing Activities, but in different order.

	Group 1	Group 2
Writing Activity 1	Topic A	Topic B
Writing Activity 2	Topic B	Topic A

The two topics will concern recent political news and recent entertainment news – popular themes in the blogosphere.

Each participant will be scheduled for a 35-minute usability testing session at a computer station. The session will be structured as follows:

- 5 minutes – Pre-test survey
- 10 minutes – Writing activity #1 without CONTEXT
- 5 minutes – Training to use CONTEXT
- 10 minutes – Writing activity #2 with CONTEXT
- 5 minutes – Post-test survey

To avoid the novelty affect, where a user may prefer a new system simply because of its new or novel features, we will randomly switch the order of Writing Activity 1 and Activity 2 for half of our participants from each of the two groups.

Sample Survey Questions

Pre-study

How often do you blog?

What are your main reasons for blogging?

What topics do you blog about?

Post-study

Which CONTEXT features did you find most/least useful? And why?

Can you tell us about actually using CONTEXT for your writing: What was easy and what was difficult?



This work is in part supported by