

InfoSMasher

Keith Bagley

CS 592: Web Systems 2 Directed Study

Fall 2008, Professor Levkowitz

Overview

The goal of the InfoSMasher project was to give users, “1-stop shopping” by gathering unified, relevant information about a selected subject matter specified. This information crosscuts a variety of mediums and sources including: text-based, video, live events, archival publications, and photographs. Originally, the InfoSMasher was targeted solely on music information mashing, as it was envisioned to be a “dashboard” for users to search the web for their favorite artist, and display the associated videos, lyrics, show dates, and artist news. However, after discussions with Prof. Levkowitz, it became increasingly obvious that mashing this type of music data was one scenario of a more general information aggregation usage model. Therefore, the solution was generalized to handle a wider range of information aggregation.

Design

InfoSMasher is implemented using Yahoo Pipes: a visual mashup programming environment. All development is performed using a graphical information flow paradigm.

The user interface consists of two major areas: the subject matter specification pane (for collecting the target subject and domain constraint from the user) and the results pane (which displays the list of aggregated information results).

The system takes as input the subject matter (or target) the user wants to capture and “mash” information about. This can be any concept, but in practice, I envision this to be either persons or organizations. The user also specifies a “domain” which further qualifies the search criteria and helps filter results. The domain can be as fine or coarse grained as the user desires. Example coarse-grained domains may be “Politics” or “Music”. Other, finer-grained domains might be “CNN” or “Boston”. The use of a domain simply helps limit the scope of matches for the result sets. InfoSMasher has some default arguments specified, so if no domain is selected by the user, “Music” is used.

Keiths InfoSMasher Web Systems 2 Project

This mashup project for WebSystems 2 (Fall 2008 with Prof Levkowitz) utilizes Yahoo Pipes to extract source data from a variety of locations, and display the associated multimedia content. It is optimized for the Music domain, but works with any domain you specify when first running the pipe.

Pipe Web Address: http://pipes.yahoo.com/pipes/pipe.info?_id=d87cb27117e17685fce6ffb7e0351e83 (edit)

The screenshot displays the InfoSMasher web application interface. At the top, there are navigation buttons: a star icon, "Edit Source", "Delete", "Unpublish", and "Clone". Below these is a "Configure this Pipe" section with two input fields: "What domain do you want to search and mash?" containing "music" and "Specify the topic or person of interest" containing "Madonna". A "Run Pipe" button is to the right. Below the configuration is a "Use this Pipe" section with social sharing options: "Get as a Badge", "MY YAHOO!", "Google", "Results by Email or Phone", and "More options". A "List" button is on the left, and "40 items" is on the right. The main content area shows a list of search results for "Madonna":

- Madonna**: Official site of pop diva Madonna, with news, music, media, and fan club.
- Madonna - MySpace**: Madonna MySpace page features news, blog, music downloads, desktops, wallpapers, and more.
- Madonna - Yahoo! Music**: Offers Madonna music downloads, photos, album information, and videos.
- Madonna Photos on Yahoo! Music**: View photographs of Madonna on Yahoo! Music. ... Madonna. Photos. Rating affects your music played in LAUNCHcast and Music Videos. ...
- Madonna - AOL Music**: Madonna official AOL Music site for Madonna music videos, Madonna songs, Madonna photos, Madonna live performances and more information on Madonna.
- 4 Minutes by Madonna, Justin Timberlake, Timbaland | Music Video**: 4 Minutes music video by Madonna, Justin Timberlake, Timbaland from the album Hard Candy
- Madonna (Entertainer) - Wikipedia**: Exhaustive bio and discography of Madonna's early life, career, "Sex" controversy, electronic club mix phase, and more.
- YouTube - Madonna - 4 Minutes (Video)**: Madonna "4 Minutes" (Video)Directed by: Jonas & FrançoisWarner Bros. Records Inc. ... Hello, you either have JavaScript turned off or an old ...
- YouTube - Madonna - Crazy For You (Video)**: MadonnaCrazy For YouWarner Bros. Records ... Hello, you either have JavaScript turned off or an old version of Adobe's Flash Player.

At the bottom, there is a navigation bar with links: "Madonna | Music Artist | Videos, News, Photos & Ringtones | MTV".

Figure 1: InfoSMasher with results

Internally, the system makes use of a number of Pipes Yahoo Search widgets to pull information from a variety of sites based on the user selection. Additionally, an automatic feed discovery widget is used to pull feed information from various sites. This aggregation of information is then filtered and duplicates are removed from the final set before being presented to the user.

Discussion

Pipes has extremely limited support for “traditional” text-based development; any development outside the graphical environment must be accessed via web services calls. Other than that, developers are effectively “locked in” to the Pipes paradigm.

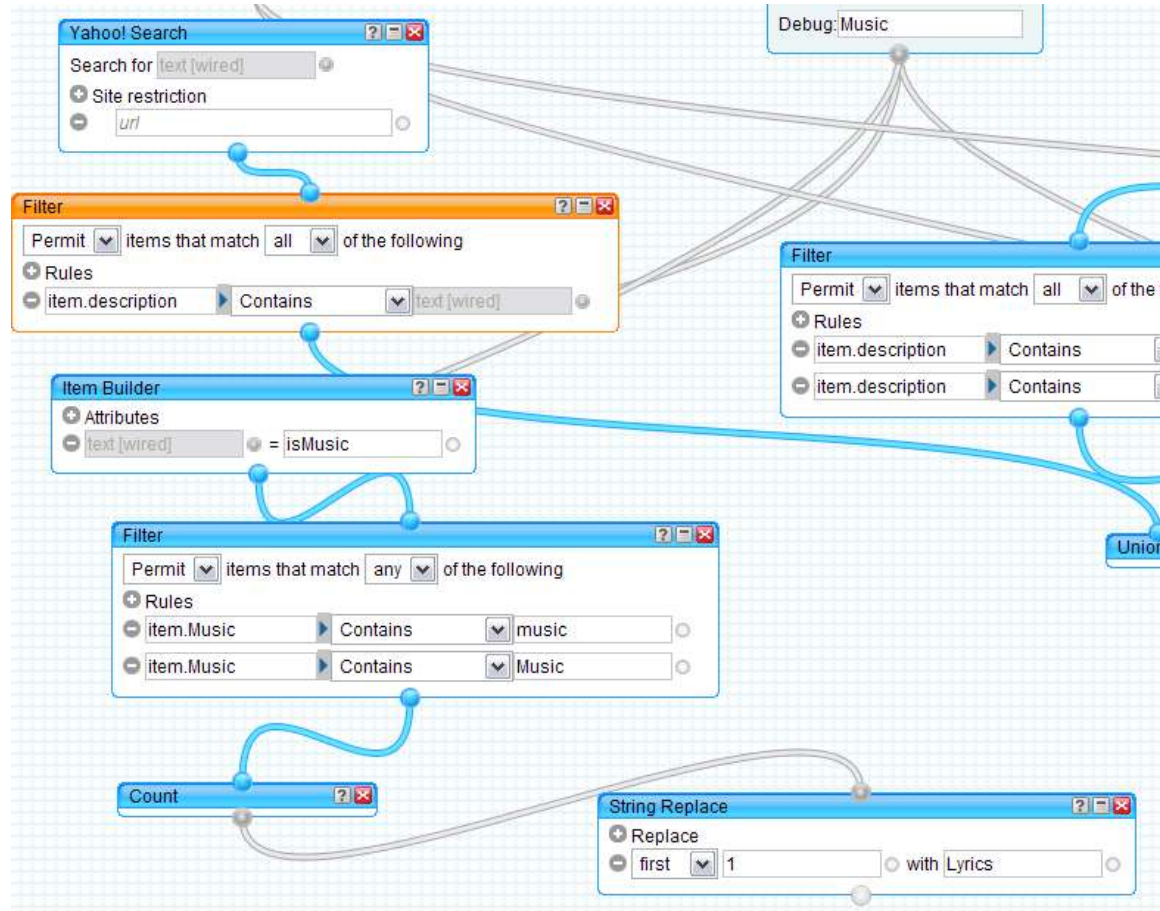


Figure 2: My Conditional Statement

The selection of Yahoo Pipes had a noticeable impact on the overall solution design. As mentioned in the overview, my initial vision of the user interface was to use a dashboard of sorts, to graphically lay out the information collected on the selected subject matter. However, Pipes is very constraining in terms of specifying UI elements and layouts. There are essentially 3 options: lists, maps, and images (in a grid). Therefore, all collected information is presented to the user in list style.

Additionally, Pipes has no notion of conditional statements. This proved extremely counterproductive for some of what I was doing, since I wanted to take one branch of actions if the user selected “Music” and another branch of actions for other domains. In the end (besides pulling out what little hair I have left on my head), I was able to simulate a simple if/then conditional statement using the Counter and Rename widgets. It was a time-consuming hack, but necessary due to the lack of common programming operators in this “language”. Figure 2 shows my solution to this problem.

In the end, it seems like Pipes is relatively easy to get started with, but as the complexity of any solution grows, the difficulty of using Pipes grows exponentially – especially if you want to do anything outside of what Pipes provides. Further, I found that the use of a graphical data-flow language only went so far. As my pipe became larger and more complex, it became extremely difficult to follow the program flow, and keep track of what was going on. This was due in part to the lack of visual abstraction mechanisms – I couldn't find an easy way to “zoom in” or “zoom out” of various abstraction levels to help me keep things straight in my head. (NB: Since I'm a Pipes Newbie, there **may** be a way, but I didn't find it). Finally, the lack of conditional statements and other common programming operators make Pipes good for a certain class of applications, but totally unsuitable for many others. This fact itself added hours of time and frustration to my project development.

Non-Pipes Development

The question: what if I had developed a similar solution using more traditional programming environments? Certainly, the infrastructure Yahoo provides saves time for many developers. I would not have been able to put together the necessary infrastructural substrate in the time allotted. Beyond that, however, the trade-offs forced on me by Pipes in terms of the UI, and in terms of having to “roll my own” conditional probably added hours and frustration on the back-end of the process. In summary, I think Pipes gives good “quickie” development for simple, single-task applications. However, for anything beyond that, I think most developers would be better suited to more traditional environments.

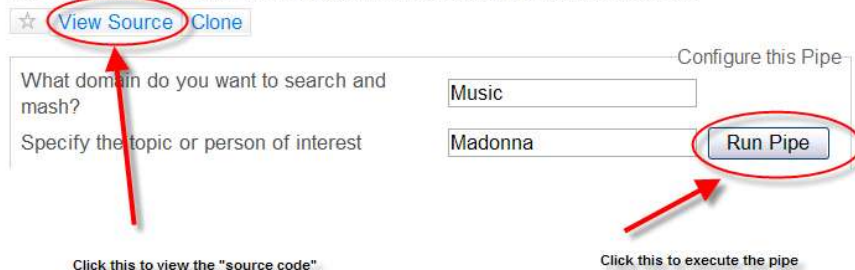
Sample Walkthrough

A sample walkthrough/demo video resides on the course wiki for those who wish to see a “live” demonstration. An abbreviated walkthrough of the application is presented here. Execute the pipe by pointing your browser to: http://pipes.yahoo.com/pipes/pipe.info?_id=d87cb27117e17685fce6ffb7e0351e83. From this location the “source code” to the pipe can also be accessed by pressing the “View Source” button.

Keiths InfoSMasher Web Systems 2 Project

This mashup project for WebSystems 2 (Fall 2008 with Prof Levkowitz) utilizes Yahoo Pipes to extract source data from a variety of locations, and display the associated multimedia content. It is optimized for the Music domain, but works with any domain you specify when first running the pipe.

Pipe Web Address: http://pipes.yahoo.com/pipes/pipe.info?_id=d87cb27117e17685fce6ffb7e0351e83



A default query is automatically run with Madonna and Music. We'll now run our own.

1. Select the domain and topic of interest. Let's try sports and Larry Bird. Sports is an example of a relatively coarse grained domain, we'll try a fine-grained one in the next example. For certain multi-word selections, it's suggested to use quotes (e.g. "Larry Bird") to perform better matches. In my testing, sometimes it helps, sometimes it doesn't matter.
2. Press "Run Pipe". The result set is displayed in the list area.

What domain do you want to search and mash? [Configure this Pipe](#)

Specify the topic or person of interest [Run Pipe](#)

Use this Pipe

Get as a Badge [MY YAHOO!](#) [Google](#) [Results by Email or Phone](#) [More options](#)

List 40 items

LarryBird.com
Browse a collection of memorabilia for the NBA great, Larry Bird. Find Bird autographs, Celtics posters, and Hall of Fame books or join the Larry Bird Fan Club.

NBA History: Larry Bird
Contains a summary of the NBA player's achievements while in the league. Featuring stats, honors, and other biographical information for Larry Bird.

Larry Bird - Wikipedia
Hyperlinked biography and background for the former NBA star, Larry Bird.

NBA.com: Larry Bird Bio
Born: 12/7/56 College: Indiana; Northwood Institute; Indiana State Drafted by: Boston Celtics, 1978 Height: 6-9 Weight: 220 ... Larry Bird was such a player. ...

Larry Bird Biography
Apparel and Merchandise from LarryBird.com including a great selection of ... He was born Larry Joe Bird on December 7, 1956, in West Baden Springs, Indiana.

Larry Bird: Biography from Answers.com
- Larry Bird 12-time All-Star, twice NBA Finals MVP and nine-time member of All-NBA First Team; led league in free-throw percentage four times

Larry Bird
TV and movie credits for Larry Bird, whose films include Space Jam and Marathon. Also features a biography and a list of notable TV appearances:

IsinTheUS Index Page
Links. Rides. 2002. 2003. 2004. 2005. 2007. William. Ice Storm ... Ice Storm Dec 2007. July 4th 2008. Easter 2008. Mother's Day. Christmas 2007. Serena & Kit ...

ESPN.com: Plain and simple, Bird one of the best
Bird gets bronzed. Larry Bird voted 30th-best athlete of century ... Larry Bird, known as "a hick from French Lick," is anything but. ...

Larry Bird Statistics - Basketball-Reference.com
You Are Here > Basketball-Reference.com > Players > B > Larry Bird. Larry Bird. Larry Joe Bird (Larry Legend, The Hick from French Lick) ...

YouTube - Larry Bird
Lawrence Joseph "Larry" Bird (born December 7, 1956) is an American retired NBA basketball player, widely considered one of the best players of all...

Figure 3: Larry Bird and Sports

- Click on a reference to open up associated content. We'll select YouTube for some video content.

The screenshot shows a YouTube video player for a video titled "Larry Bird". The video content is a CD-ROM with a portrait of Larry Bird and the text "SMALL TOWN performed by John Mellencamp". The player interface includes a progress bar at 0:04 / 3:14. Below the player, there are social sharing options for MySpace, Facebook, and Digg, along with a "Send this video as a greeting card" button. The video has a rating of 5 stars (1 rating) and 340 views. To the right, there is a sidebar with the channel "joostvdenden" (joined July 10, 2008), a "Subscribe" button, and a list of "Related Videos" including "Larry Bird - Larry Legend", "Larry Bird: Small Town", "1991 NBA Playoffs: Larry Bird Comeback", "LARRY BIRD SAYS N-WORD", and "Larry Bird". At the bottom right, there is a "Promoted Videos" section with four video thumbnails.

Figure 4: Multimedia content from our Mashup

- Close the YouTube window (or tab). Perform another mash. This time, select *cnn* as the domain (this is relatively fine-grained), and Wolf Blitzer as the topic. The result set again appears.

Keiths InfoSMasher Web Systems 2 Project

This mashup project for WebSystems 2 (Fall 2008 with Prof Levkowitz) utilizes Yahoo Pipes to extract source data from a variety of locations, and display the associated multimedia content. It is optimized for the Music domain, but works with any domain you specify when first running the pipe.

Pipe Web Address: http://pipes.yahoo.com/pipes/pipe.info?_id=d87cb27117e17685fce6ffb7e0351e83 (edit)

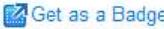




☆ [Edit Source](#) [Delete](#) [Unpublish](#) [Clone](#)

Configure this Pipe

What domain do you want to search and mash?

Specify the topic or person of interest

Use this Pipe

List 19 items

Show Pages - The Situation Room - CNN.com
 Nightly newscast that focuses on the day's top news, live interviews, and debriefs with CNN correspondents.

CNN Programs - Anchors/Reporters - Wolf Blitzer
 Wolf Blitzer is the anchor of CNN's The Situation Room, CNN's innovative, fast ... Blitzer also hosts Late Edition with Wolf Blitzer, the only Sunday talk show ...

Wolf Blitzer: News & Videos about Wolf Blitzer - CNN.com
 Alaska Gov. Sarah Palin, this year's Republican vice presidential nominee, spoke with CNN's Wolf Blitzer on "The Situation Room" on Wednesday.

Wolf Blitzer | NewsBusters.org
 When both Borger and host Wolf Blitzer both affirmed that he did have ties to ... CNN's Wolf Blitzer Interrupts Serious Election Discussion for The Boss' ...

CNN Political Ticker: All politics, all the time Wolf Blitzer " - Blogs ...
 From CNN Anchor Wolf Blitzer. What impact will Nader and Barr have this election season? ... CNN Anchor Wolf Blitzer. Watch CNN's Wolf Blitzer and David ...

Wolf Blitzer for the Defense (Department)
 On the rare occasion when a mainstream news program interviews a forthright critic of U.S. policy, the interviewer often ... when CNN's Wolf Blitzer (11/7/02) ...

Wolf Blitzer
 What is the One Question Wolf Blitzer Can Ask Obama He Hasn't Answered Already? ... Read More: Portman McCain, Wolf Blitzer, Rob Portman, McCain Sec Chair, Cnn Late ...

Wolf Blitzer - wtop.com
 ... of Presidential Race. Wolf Blitzer, CNN. Posted Online: 10 ... Wolf Blitzer, host of CNN's Late Edition and The Situation Room. Posted Online: 09-19-08 17:09 ...

Wolf Blitzer - Wikipedia, the free encyclopedia
 Wolf Blitzer (born March 22, 1948 in Buffalo, New York) is ... "Wolf Blitzer for the Defense (Department)", Jim Naureckas, FAIR Extra!, January/February 2003 ...

Figure 5: Wolf Blitzer and CNN

Conclusion

This was an interesting project that, although small in scale, I believe has merit for future research and application. Mashups are the buzzword du jour, and beyond being fun for technophiles to produce, there is definitely potential for very useful information aggregation while using the web as a general purpose repository. On the other hand, in my opinion, the use of Yahoo Pipes seriously limits the production of large-scale solutions of this sort. Other mashup editors and tools (IBM Lotus Mashups, Google Mashup Editor, Intel Mashup Maker) may all offer a better developer experience and more “enterprise level” support for creating larger and more complex solutions. So, while I think the mashup concept is useful and ready for primetime; I believe the graphical paradigm of Yahoo Pipes needs a bit more refinement. At the very least, opening up the system to allow for easier non-graphical development (when necessary) would be a good starting point.

For future work, I believe that a customizable, hybrid approach to creating mashups of this sort (allowing for graphical development when desired and text-based traditional coding when needed) would move mashup production to the next level of evolution and adoption.