Taxonomies for News Aggregator Interfaces and User Patterns

Filipe Fortes
Center for Interaction Design
Microsoft Research Asia
Beijing, China
fil@fortes.com

Ryan Lv
Shanghai Jiaotong University
Shanghai, China
hlv@acm.org

David Dehghan
Microsoft Corporation
Redmond, WA
davdeh@microsoft.com

Neema Moraveji
Center for Interaction Design
Microsoft Research Asia
Beijing, China
neemam@microsoft.com

Abstract

In this paper we define terminology for discussing the design and usage patterns for news aggregator applications. The work rests on a thorough analysis of available news aggregators and contextual inquiries with 34 users. The work explains approaches to aggregator user interface design and how those interfaces are employed. The results motivate the design of aggregators addressing these patterns.

Keywords

News aggregators, RSS, reading, information retrieval, weblogs

ACM Classification Keywords

H5.4. Information interfaces and presentation (e.g., HCI): Hypertext / Hypermedia.

Introduction

Information proliferation has recently begged the design of *news aggregators*, applications that allow a user to stay up to date with many web sites without visiting each site individually. These applications have dramatically increased the efficiency of navigating information sources and have led to the evolution of new interface designs and use strategies.

While working on the creation of an aggregator application, we performed a thorough analysis of existing aggregators. Additionally, we performed thirty-four user interviews, where we observed aggregator use.

In this paper, we present the products of our investigations: terminology for discussing the differences between common aggregators and the usage patterns observed during user interviews. These findings can aid us and the research community in designing and discussing aggregators.

Background

An *aggregator* is software that collects syndicated web-based information from multiple sources (*feeds*) within a unified interface. A feed (commonly called RSS feed or news feed) is a document typically stored in an XML-based format, the most popular being RSS (Really Simple Syndication / Rich Site Summary) and Atom.

Each feed contains a set of *entries*, the content of which can vary dramatically. In summary-only feeds, entries contain a short textual abstract with a link to a full version of the content, while the entries in full text feeds include the full content within the entry. Optionally, an entry may have an *enclosure*, a reference to file to be included with the entry, much like an email attachment.

Originally, feeds were used exclusively for distributing news headlines. Thanks to the flexibility of the format, today it is common to see feeds being used for weather forecasts, and personal calendar data, and multimedia such as podcasts (audio/video enclosures).

The diversity of feed content has lead to a similar diversity in applications that can consume web feeds. For example, Apple Computer's iTunes, a popular digital music jukebox software package, has built-in support for podcast feeds. The audio or video enclosures are automatically downloaded and added to the user's collection. Users cannot subscribe to non-podcast feeds within iTunes, because the specialized software only handles feeds with audio/video content.

We concentrated on news aggregators, which are primarily used to collect and display text-heavy content. Even within this niche, there is large degree of variety within applications.

Aggregator Terminology and Taxonomy

The user interface of a news aggregator can vary dramatically. In order to frame the space of these applications, we must first develop a standard vocabulary for the salient characteristics of these applications. We present two taxonomies for classification developed by analyzing the most popular choices on the market.

The first taxonomy describes the environment in which the aggregator software is executed:

- *Desktop aggregators* are client-side software that run on a user's machine and are displayed in their own desktop window.
- *Web-based aggregators* run on a remote web server and are accessed through a web browser.
- *Widget aggregators* typically occupy a small portion of screen real-estate and are integrated into a web
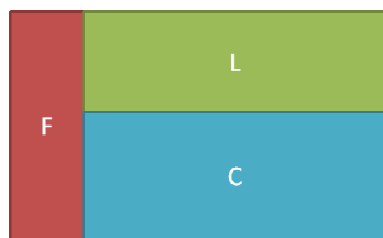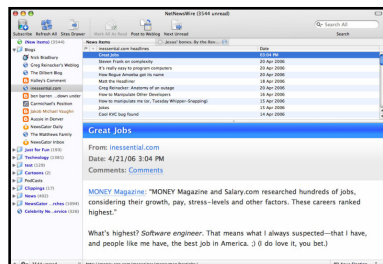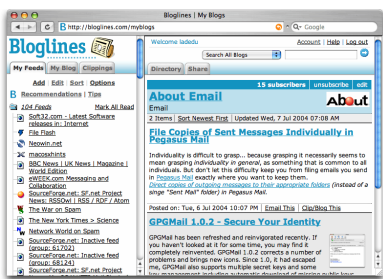
Figure 1 (above): NetNewsWire, a popular desktop aggregator, displays all three panes with classic three-pane interface.

Figure 2 (below): Bloglines, a popular web-based aggregator, shows only two panes (F and C).

portal page such as NetVibes or desktop sidebar such as Yahoo!'s Konfabulator.

- *Mobile aggregators* run on smaller devices such as mobile phones or PDAs.

Some aggregators provide multiple clients; for example, NewsGator provides both desktop and web-based clients.

The second taxonomy describes the user interface of an aggregator, as well as the actions that interface affords. At first glance, this taxonomy is not straightforward, so we will be begin by using an example which is familiar to most users – the user interface of a typical desktop email client such as Microsoft Outlook Express or Mozilla Thunderbird.

Modern desktop email programs display three panes within the application window: The folder (F) pane lists the user's email folders and is typically displayed on the left side. The list (L) pane displays messages within the folder selected within pane F; the L pane is typically displayed on the top half of the screen. The content (C) pane displays the content of the email selected in pane L; the C pane occupies the remainder of the screen, typically the lower half.

Many aggregators are based on email or newsgroup clients, so we use the same terms to describe their parts:

- *Folder (F) pane*: Used to select the source of the entries being displayed. Like email clients, most aggregators allow users to organize their feeds into

folders or categories. Additionally, many aggregators provide virtual or search folders that display entries that meet a set of criteria (e.g. unread entries whose title contains "banana"). Not all aggregators support feed categorization, nor do all aggregators display this pane.

- *List (L) pane*: Used to select an individual entry to be displayed. Unlike mail clients, many aggregators can display of multiple entries at a single time – some of these aggregators do not provide this pane at all.

- *Content (C) pane*: Used to display the content of an entry or entries. Unlike mail clients, the amount of information displayed per entry can vary from only the title of the entry, to the full content of the entry, to the website URL associated with that entry. Without exception, all aggregators have this pane.

By analyzing the most popular aggregators on the market, we found distinct groupings of interfaces:

- *Three-pane / Mail-client aggregators* provide interfaces are nearly identical to email clients. Like mail clients, the content pane displays only one entry at a time. All three panes (F, L, C) are available.

- *River of News aggregators* mix the entries from all of the user's subscriptions into a single view. Entries are presented in reverse chronological order, and frequently no distinction is made between read and unread entries. Only one pane (C) is displayed, the other panes (F, L) do not exist.

Figure 3: Classifications for selected aggregators.

Aggregator	Env.	UI
My Yahoo!	Web	Headline (C)
Bloglines	Web	Entry river (FC)
NetNewsWire	Desktop	Mail client (FLC)
NewsFire	Desktop	List river (LC)
Radio Userland	Desktop	River (C)

Figure 4 (below): Yahoo!'s page uses Headlines.
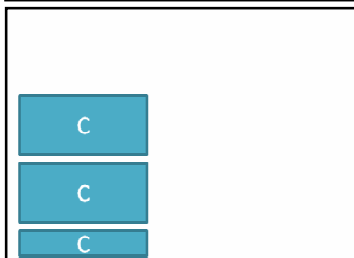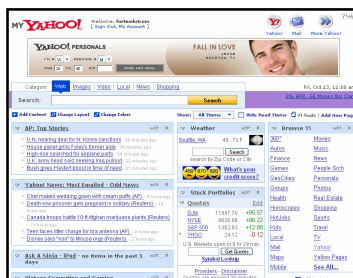
- *Headline aggregators* display multiple entries within a compact space. Typically only the title (headline) of the entry is displayed, although some provide a one-line summary excerpt as well. This style is almost exclusively used by widget-based aggregators and other instances when there is very little space. This display is considered a content (C) pane.
- *Two-pane aggregators* display characteristics of both three-pane and river of news aggregators. There are two variations, based on which two panes are displayed:
  - *List River (L, C)*: The list pane displays a river of entries across the user's subscriptions, while the content pane displays only an entry at a time.
  - *Entry River (F, C)*: The content pane displays river entries from the source selected in the folder pane. Unlike the standard River of News, entries are typically grouped by feed before being sorted.

Note that we have described each of these interfaces in their default configuration – many aggregators allow the user to show or hide pane via configuration or keyboard shortcut, which may modify the interface slightly. Additionally, some aggregators provide slight hybridizations between styles[1].

---

[1] In RSS Bandit, a popular Windows desktop aggregator, the content pane will display multiple entries like an entry river unless the user clicks on an entry within the feed pane – at which point only that entry will be shown in the content pane.

Observed User Behaviors

Thirty-four user interviews were conducted in order to observe patterns of behavior across aggregators. Interviews lasted approximately 60 minutes each and were held in the participants' primary aggregator use location (home or office). Participants were instructed to perform a cognitive walkthrough as they went through a live feed reading session.

Interviewees included software engineers, students, executives, and journalists. All but two of the participants were male, and all had several years of computer experience. Thirteen participants were found by inspecting the list of users subscribed to a local weather feed on the popular web-based aggregator Bloglines. The remaining twenty-one were found through a process of informal referral.

Over two-thirds (24) of the users we observed use an entry river (FC) aggregator; this is undoubtedly influenced by our recruiting techniques, since all users of Bloglines (18) fall into this category. Of the remaining users, 4 used a mail-client (FLC), 3 use an list river (LC), one uses river of news (C), and one used a headline (C).

Several patterns in user behavior emerged from cross-user analysis; specifically, we found interesting patterns for

- Folder/Feed Selection
- External Navigation Technique
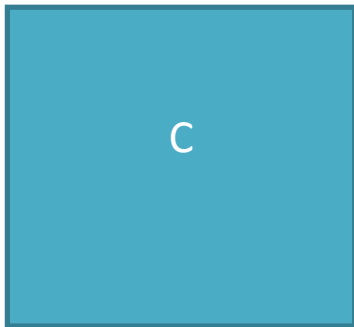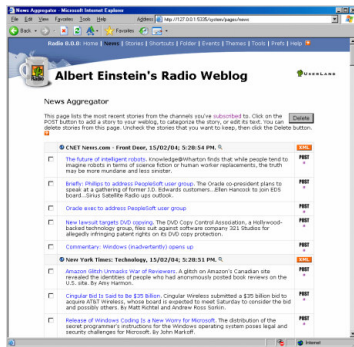- Session Termination and unread items

Figure 5: A web-based aggregator employing River of News.

Folder/Feed Selection

The overwhelming majority (29) of users interviewed used an aggregator that displays a folder (F) pane. Amongst those users, we observed three strategies for selecting which folder or feed to read from the pane:

▪ *Routine*: Slightly over half (15) of interviewees reporting having a specific routine selecting feeds. These users reported having a set of feeds that they always read, frequently in a specific order.

▪ *Sequential*: About a quarter (7) of interviewees traversed through folders and feeds in a sequential order. Although these feeds were typically organized into categories, the users did not say the order of the categories was important; in fact, one user started from the bottom and moved upwards.

These users occasionally skip over a feed that they did not want to read at that time. In a few cases this was because that feed had too many in unread items. In two cases, the feed was in a language the user was not fluent in, and the users reported not wanting to switch languages.

▪ *Whimsical*: All but one of the remaining users (n=6) scanned through their feed list picking feeds seemingly at random. Two users reported searching for feeds that were quick to read, either with few unread items or feeds they knew to have content that was easy to consume.

The final user utilized a hybrid strategy, first reading a fixed set of feeds (routine), followed by a sequential reading of the remainder.

External Navigation Technique

Users frequently clicked on hyperlinks while reading, with all following at least one link during the interview. Users clicked on links for one of three motivations:

▪ *Read the full entry*: If the full content of an entry is not displayed, then the user must follow a link in order to fully read the article..

▪ *View / leave comments*: In order to read or leave comments on the entry.

▪ *Follow a reference*: Following a link referenced within the entry.

Across all three motivations, we observed two user strategies when following a hyperlink:

▪ *Mark and Sweep*[2] half (n=17) of users open the link (mark) and continue reading entries, letting the page load in the background (typically in a new tab in their web browser). Users read (sweep) the links at a later time – a third of these users (n=7) swept once they had finished their current feed, another third (n=6) swept only after they were done with their reading session.

The remaining users didn't follow a consistent pattern, although exhibited interesting behavior: One user waited only as long as it took for the first of the marked articles to finished loading in the background. The other user would begin a sweep whenever the

---

[2] These names are analogous to the names of two well-known algorithms for garbage collection: *Mark and Sweep* and *Stop and Copy*.

aggregator was slow to respond, which happened when switching to a feed that had many unread items.

▪ *Stop and Read* just under half (n=15) of users suspended their aggregator use in order to immediately read the link they had opened.

The remaining two users used both strategies within the interview, opening some links in the background, and reading others immediately.

Session termination and unread items
All interviewees had over 100 unread items at the start of their reading session – one interviewee had roughly 20,000 unread items! Only three users used aggregators that do not display read/unread information for entries. Three other users ended their session with no unread items, while the remainder (n=28) all left over 100 items unread.

Just under one third (10) of users reported not caring about the read/unread status of their entries. Five users specifically reported that the read/unread flag was important to them. Two of these users reported they used the manually kept items as unread as a reminder.

Conclusion

This work first defined the category of news aggregator applications as it applies to the user interface design community. This was followed by an analysis of 34 user interviews resulting in several notable user behaviors. These behaviors, when combined with the terminology of aggregator interface design, can serve as a base of discussion and motivation of aggregators in the community.

References
[1]   Denning, P. (1982). Electronic Junk. Communications of the ACM, 25(3), March 1982.

[2]   Bloglines. http://www.bloglines.com/.

[3]   Hornbæk, K. and Frøkjær, E. 2001. Reading of electronic documents: the usability of linear, fisheye, and overview+detail interfaces. In *Proceedings of SIGCHI.* CHI '01. ACM Press, New York, NY.

[4]   Kidd, A. 1994. The marks are on the knowledge worker. In *Proceedings of SIGCHI.* CHI '94. ACM Press, New York, NY.

[5]   Mackay, W. (1988). Diversity in the Use of Electronic Mail: A Preliminary Inquiry. ACM Transactions on Office Information Systems, Oct. 1988.

[6]   Pirolli, P. and Card, S. 1995. Information foraging in information access environments. In *Proceedings of SIGCHI*. CHI '95. ACM Press, New York, NY, 51-58.

[7]   Venolia, G., Dabbish, L., Cadiz, JJ, Gupta, A. Supporting Email Workflow, MSR-TR-2001-88.

[8]   Whittaker, S. and Sidner, C. 1996. Email overload: exploring personal information management of email. In *Proceedings of SIGCHI.* CHI '96. ACM Press, New York, NY.