

What Do We "Mashup" When We Make Mashups?

Jeffrey Wong

Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, PA 15217
jeffwong@cmu.edu

Jason Hong

Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, PA 15217
jasonh@cs.cmu.edu

ABSTRACT

This paper describes preliminary work in the uncovering of mashup patterns in order to find new directions for the design of mashup tools. We conducted a qualitative survey of high-quality mashups, as nominated in two popular mashup directories, and examined how the mashups made use of existing websites or improved upon them, how data from multiple websites were combined, and what kinds of user tasks these mashups might be suitable for. We describe a set of patterns what we found in our sample of mashups.

Categories and Subject Descriptors

H.5.2 [User Interfaces]: Style guides; D.2.13 [Reusable Software]: Reuse models.

General Terms

Design, Human Factors

Keywords

Mashups, Patterns, Taxonomy

1. INTRODUCTION

Much prior work on mashups has described mashups as websites that combine the functionality of two or more websites. But mashups are actually more than that. Mashups are inspiring and interesting because they are an expression of how users make use of existing systems in new and innovative ways.

Mashups and tools to support the construction of mashups are important to the development end-user programming [6]. Use of the web is pervasive for many tasks that involve searching for information or decision-making. Constructing mashups typically involves programming, although there are now a number of tools that simplify or eliminate programming for a number of mashup tasks.

Mashup construction tools are often designed to address different mashup patterns. Microsoft Popfly [5] and Marmite [8] conceive of mashups as data flows that take data originating from a source or user input, applying web services to augment or transform the data, and visualizing the results on a map or on a web page. Yahoo Pipes [9] takes data from RSS feeds or web services and applies

programming operations such as text filtering, counting, sorting, and regular expressions. Karma [7], a data-integration by-demonstration system focuses on data unification where disparate data sources are linked together by common attributes, similar to "joining" database tables. Enterprise mashup tools, such as IBM's QEDWiki [4] focus on connecting graphical visualizations to data sources, following a pattern established by the first mashup, HousingMaps (housingmaps.com).

With tools following very different models of what a mashup is, what is meant by the term "mashup" becomes unclear. Mashups originally referred to websites combining features of multiple websites. Some mashup tools enable users to construct entire websites as a final artifact [5]. However, that definition has been blurred by tools that focus on data integration issues (such as Karma [7]). Also, adaptation of existing websites to alternative environments like mobile devices is the goal of a number of mashups listed in the mashups directory website ProgrammableWeb.com, as well as a design goal of the d.mix [2] web service integration tool.

The definition of mashup is also somewhat dynamic. There is an entire category of mashups known as "map" mashups. Most likely owing to the inspiration from HousingMaps, these mashups comprised over half of the mashups listed on ProgrammableWeb from 2005 to 2007. Many of these were simply a static data set or results from a query visualized on a map. In April of 2007, Google introduced Google MyMaps, a feature of their mapping tool that enables users to plot customized data on a map without having to learn the Google Maps API and create a website of their own. Subsequently, as of January 2008, maps only comprise 39% of all mashups listed on ProgrammableWeb. Perhaps this is evidence that many map mashups are not mashups at all, but simply uses of maps as visualization widgets, like charts, tables, and selection menus.

Since mashups appear to follow various kinds of patterns and each mashup tool aims to support specific patterns, we propose that it is valuable to survey the space of mashups and catalogue what those patterns are. The idea of patterns has been used in the past to present generalized solutions to problems that may be encountered when designing websites [3] or writing software[1]. Understanding patterns in mashups can help drive the development of mashup tools to support specific patterns. To our knowledge, there has been no survey of mashup patterns. In the rest of this paper, we describe a preliminary survey we conducted and report on observed patterns.

2. SURVEY METHODOLOGY

To come up with some mashup patterns, we qualitatively surveyed a small set of mashups from two mashup directories that showcase mashups. For each mashup, we looked at what APIs were used (according to what the each mashup author claims in the

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

WEUSE IV'08, May 12, 2008, Leipzig, Germany.

Copyright 2008 ACM 978-1-60558-034-0/08/05...\$5.00.

directory), what data sources were involved, and what websites each mashup used. We also attempted some guesses at what purpose the mashup might serve or how each mashup improved on existing websites either functionally or in terms of user experience.

We also assigned category names to describe these improvements. For example, oSkope (www.oskope.com) provides an innovative visual search interface to Amazon and eBay that doesn't require users to visit multiple pages. It also enables users to build up ad-hoc lists of items of interest, which is more difficult on Amazon and eBay where items need to be saved to a "wishlist" triggering a page reload. Surprisingly, although the mashup makes use of multiple websites, it does not unify data or integrate searches across them. We assigned this mashup to the "Alternate UI" category.

We examined high-level user inputs and outputs. For example, the "last.fm normalizer" accepts a last.fm user name and shows statistics about that user's playlist. We attempted to guess at the underlying information structure such as how data is unified between multiple websites or otherwise transformed and presented to the users. We also came up with binary labels to describe interesting features in each mashup. These labels were created whenever we noticed a unique way that a mashup improved upon the original website and then checked to see if any mashups we had already examined would deserve the label. Table 1 explains some of these labels.

Table 1. Some binary categories for each mashup

Search	Is the mashup a search interface?
Visualization	Does the mashup add visualization to the data?
Real-time	Is the purpose of the mashup to allow the user to monitor or observe the original website as a real-time data set?
Widget	Is the mashup actually a widget for some platform like Apple's Dashboard, the iPhone, or Google Homepage?
Personalized	Does the mashup make use of the user's personal information from the original website or enable the construction of a personalized data set from the original website?
Folksonomy	Does the mashup make use of a tagging system or adds tagging to the original data set?
In-situ use	Is the mashup simply a tailoring of an original website optimized to a specific situation use?

It would be difficult to apply this qualitative analysis to all mashups in existence. On ProgrammableWeb, there are over 2700 mashups in its mashup directory. Part of this is due to the fact that mashup authors submit their own mashups to the ProgrammableWeb directory. However, many of these mashups are poorly thought out, or assembled only for demonstration purposes and submitted to the directory to collect community

feedback or to drive advertising revenue by being linked from a highly reputable site. In order to make this study feasible, we examined a random sampling of mashups. To reduce the number of bad mashups sampled, we took a random sample of 22 mashups that were in the set "Mashup of the Day" on Programmable Web. We also examined mashups from MashupAwards.com, which accepts nominations of interesting websites and selects one mashup each day to be given a "Mashup of the Day" award.

End-user customization of the web experience is not just limited to mashups. We also looked briefly at a sampling of popular GreaseMonkey scripts. Some Firefox extensions might also be classified as end-user customization of websites but we did not examine these.

Furthermore, we do not claim our methodology is completely objective although we believe it will yield valuable insights.

3. SURVEY FINDINGS

Our findings are Table 2, where ad-hoc categories of mashups as well as the labels can be taken as mashup patterns. There are several categories and labels we found were interesting.

Aggregation. A common function of mashups is to aggregate multiple websites together or summarize sets of data. But this takes on multiple forms. EveryBlock takes an address and find nearby geotagged pictures from Flickr, finds local restaurants, finds reports of nearby crimes and other public notices, and looks for nearby craigslist ads. This is similar to entering a single location as query into multiple websites and assembling the results onto a single page. soup-soup is similar but uses keywords from news stories instead of single query terms to search wikipedia and Flickr for relevant information. Vidmeter finds the most popular videos from multiple websites and assembles them into a single ranking chart. This probably requires some sorting and addition. Trendite is a site that supports monitoring of search trends. It most likely makes requests for top search queries and results from the major search engines and assembles them onto a single page. Song List assembles links to song samples and needs to unify song titles between multiple music services.

Alternate UI & In-situ Use. These mashups don't combine multiple websites at all but rather aim to support new methods of interacting with data from the website or support specific use cases. Leaflets (getleaflets.com) are specialized versions of common websites, such as upcoming.com, that are optimized to run on the Apple iPhone, thus saving users from having to visit the normal versions of those websites. Tagbrowsr provides an alternate form of tag-based browsing to Flickr's search-by-tag function. oSkope supports searching and browsing better than Amazon and eBay's normal search tools by reducing page loads and support ad-hoc collections.

In-situ use mashups support the specialized use of a website outside of the typical use case. HeyWhatsThat compiles terrain elevation data to produce printable drawings that can help a person determine what landmarks are visible from a vista point. delimport is a plug-in for the del.icio.us bookmarking service that enables Mac users to access their bookmarks through a keystroke-based search facility built into the operating system, instead of visiting the del.icio.us website.

As expected, popular GreaseMonkey scripts found on userscripts.org tended to be modifications to the user interface of sites. Aggregation by unifying with data from another web site was rare.

Table 2. Selected Mashups and Patterns (See Categories and Labels Columns)

Mashup	APIs	Notes	Category	Labels	Inputs	Outputs	User questions or needs addressed	Information Structure
delimport ianhenderson.org/ delimport.html	delicious	Enable Spotlight on MacOS to index your delicious bookmarks	Alternate UI	Personalized, In-situ use, Search	your delicious id		I want keystroke access to my bookmarks anywhere on my computer	Grab list from delicious, put in Spotlight
EveryBlock www.everyblock.com	Yelp, craigslist, flickr, City information feeds	Find out what is happening in your neighborhood including official city business info. (could lend a lot more usefulness to the intent of public postings in newspapers)	News, Aggregation, Personalized Filtering	Real-time, Personalized	Feeds from a variety of different data sources	Feeds of various kinds filtered to relevance about someone's neighborhood	What's going on my neighborhood	Addresses are used to filter events
ExitAhead www.exitahead.co.uk	last.fm, eBay	like the last.fm normalizer. But how hard is it to match up artist names with items on eBay? Is this a really difficult problem where the best	personalized shopping with personal data from a lifestyle site	Search	<ul style="list-style-type: none"> •your last.fm name or friends last.fm name •settings about how to search eBay •what part of your list do you want to use 	A list of links to expiring auctions on eBay from a person's, RSS Feed	Enter mass search inputs based on personal interests	feed of search results, network links to social network
FlickrRealTime www.pimpampum.net/rt/	Flickr		see changes in realtime	Real-time				
HeyWhatsThat www.heywhatsthat.com	GoogleMaps, Google Earth, USGS Elevation data	Complex, featureful vista data. Print out of what your view will be at a particular summit	Visualization	Search, Visualization, In-situ use	Selected vista	Interactive visualization of nearby peaks	What can I see from where	Radar topography data mixed with list of peaks from USGS
last.fm normalizer www.associativetrails.com/stuff/ normalisefm/ index.cfm	last.fm	Reveals statistics without passwords?	Compute statistics, Reveal patterns not normally visible.		last.fm username, selection about output wanted		Show underlying data not normally present in an existing interface. Compile information in a new way	
Leaflets getleaflets.com	Various	Use alternative interfaces to various websites that are optimized for iPhone	Alternate UI	Widget			I want to access certain services without visiting the actual website	
oSkope www.oskope.com	Amazon, eBay,	not unifying data, provides details and links to original pages, support ad-hoc exploration, still need to customize by bringing necessary details up	Richer interaction, Alternate UI	Search, Visualization	search terms or need	choices, links to original pages	<ul style="list-style-type: none"> •I want a movie/ used thing/book/ video/picture and I want to pick and sort by various criteria, •Save me from stupid page navigation •Help me consider things 	Lists, axes
PackMapr packmapr.com	UPS, USPS, FedEx, Google Maps	Enter a tracking number on any service, see path traveled, create an RSS feed that will be updated when its path changes	Simplification, Notification	Visualization, Real-time	Tracking number from any service	Map with data, RSS feed	Where's my stuff? Where the heck are these towns that it's going to	Tracking number leads to list of locations
Song List www.song-list.net	iTunes, Napster, Rhapsody,	Just a listing of artist names. Aggregates artist names on 3 different music websites.	Aggregation	Search	artist name, song selection	links to song sample on affiliate website	Find me a sample of a song from an artist but I don't want to have to go to each website to see if they have it.	Unification of artist and song records between multiple music websites
soup-soup www.soup-soup.net	News sites, Flickr, wikipedia, & Technorati?	Each news headline is appended with somewhat related Flickr photos and wikipedia articles. Good for getting the background on news topics. Wouldn't you have to read the article first to have the desire to look something up though? A browser plug-in might make more sense.	News, Augmentation		News headlines	Pictures and Wikipedia links	I want to read the news but I will probably need more background info	(not sure how they get the summary terms out of the articles)
TagBrowsr www.tagbrowsr.com	Flickr	Sort of faceted browsing style of selecting sets of tags. Navigate along tag groups, sets, pools, and photographer	Alternate UI	Search, Folksonomy	initial tag + tags to add to the set			multi-network of photos where each photo is a node in a variety of networks, adjusting the tag set, changes what network you are in
The Community Broadband Map www.ecorridors.vt.edu/maps/ broadbandmap.php	Google Maps	User-generated measures of broadband speed tested and then placed on a map	Visualization, User-generated content	Visualization	Map view selection			Private database visualized
TreeHugger Green Index www.treehugger.com/grndx.php	Daylife	Tracks share of mentions of environmental topics in the media. This mashup supports on a larger website rather than being on its own. Unclear what the index means in the real world.	News, Aggregation, Monitoring, Stats		Green terms	Media stats (table)?	What are the current media trends with respect to environmentalism?	Data point in media are aggregated into numbers

Table 2. Selected Mashups and Patterns (See Categories and Labels Columns)

Mashup	APIs	Notes	Category	Labels	Inputs	Outputs	User questions or needs addressed	Information Structure
Trendite www.trendite.com	Various search engines	Monitor what's hot in searching	Monitoring, Aggregation	Search	Search engine feeds	Report of what's hot and moving	What are people searching for?	
TwitterWhere twitterwhere.mattking.org	Twitter	Generate feed for all twitters in your area	Feed, Personalized	Real-time	Location	Feed of Twitters	What are people in my area thinking?	Filtering of stuff from twitter
Vidmeter www.vidmeter.com	YouTube, MySpace, Google Video	Show most popular videos across major video viewing sites.	Aggregation, Summarization	Search, Real-time	selected rankings settings	List of popular videos	I want to kill time and see what's hot	Ratings and viewing stats from multiple video sites are compared for daily breakdown
VisualSearch (Ideé) labs.ideeinc.com/visual	Alamy.com	Fast and simple. Matches against semantic and visual similarity. Faster than using Flickr. I could have used this when hunting for images	Visual Search, Non-specific search, Novel Algorithm	Search	Tag or initial image selection	Links to stock image DB	•I need some stock imagery that fits my color palette and has some meaning	Network, tags
Wishpot www.wishpot.com	Amazon (for importing wishlists mostly)	Helps users aggregate their wishlists into one place. Browser bookmarklet enables capture from any website. Widgets for mySpace and facebook let you make your desires part of your identity. It's a tool that takes snippings from multiple websites, helps you buy those things	Personal Aggregation, Clippings, In-situ use	Personalization, In-situ use	Clippings from other websites, SMS ISBN numbers or photos, Amazon wishlist	A list, price comparison info, mySpace and Facebook widgets	I want to keep track of everything I've wanted and get people to buy stuff for me if they see this info on facebook.	products can be referenced by barcode or probably screen-scraping of URLs that people submit
YES for eBay yes.sagefire.com	eBay	Computes a year-end summary of your sales performance on eBay. eBay 2007 developer award. This is a service that eBay could have thought of but an outside individual created. Probably uses the complex eBay API in ways that are unforeseen.	Stats, Data Analysis, Summarization	Personalized	eBay user name	YES report	Tell me how I did on eBay	N/A, have to look at API for how YES data might be computed
Yoututorials yoututorials.com	Youtube	User-submitted categorization of tutorial videos found on Youtube. Moderator accepts and categorizes videos. Adds tags to Youtube videos. Tags classify tutorials however.	Subset View of Data	Folksonomy	tag or category name	list of videos	What Youtube tutorials are there?	Videos have tags and are in a category

Personalization. A number of mashups personalize based on either personal information from the websites they are based on or new personal information from users. YES for eBay computes a year-end summary of sales and tax liabilities for a person's eBay account. Wishpot uses browser scripts to create personal wishlists of items a user is interested in buying in the future as they encounter them on other websites. ExitAhead takes an artist list from a user's last.fm playlist and compiles links to related paraphernalia.

Focused View of Data. This pattern is where a mashup exists to index or categorize a subset of another website's entire contents. Yoututorials is a user-submitted catalog of tutorials that can be found on YouTube. Tags and a category system are further applied onto the set of videos to facilitate searching.

Real-time Monitoring. A number of mashups support real-time monitoring. We do not mean real-time in the systems sense but rather mashups that make a changing data set the focus of the design and are intended to make the user aware of changes. TwitterWhere and FlickrRealTime show updates every few seconds, while Vidmeter might change only a few times an hour. EveryBlock's real-time data comes from pictures, added restaurant reviews, and public notices, which are all likely to see change on the scale of days.

4. DISCUSSION & FUTURE WORK

We have presented a few interesting patterns that we have observed so far in this preliminary study. It is possible that there are more patterns in the mashups we have encountered so far. However the diversity of the forms of the mashups in our data set requires us to examine a larger set in order to see the patterns for ourselves. Single data points do not make patterns.

We noticed that single source mashups tended to fit the Alternate UI or Real-time Monitoring pattern. Although most mashup tools do not support the construction of complex user interfaces without dealing with HTML, there has been past work in building user interfaces without programming that might be applied to this situation. Apple's Safari web browser currently supports creating Dashboard widgets by selecting a DOM element from a web page that contains the data that will be of future interest. We expect to see more mashups which are specialization interfaces.

Our patterns come from high-level guesses of how mashups make use of other websites and web service APIs. Mashup patterns may become richer when we look closely at the details of how data from other web sites and services are used. Ideally, the ground truth in mashup patterns can be found in the code itself, but it is unlikely that a survey looking at the source code of many mashups will be feasible. Future work may examine the APIs and data that mashups use more closely, reverse-engineering the mashups in a sense.

In this study, we only looked at mashups that are whole websites. However, we have yet to look at mashups or other artifacts that result from mashup tools. These instances may not be complete mashups in and of themselves but they still represent end-users efforts towards customizing their web experiences.

One caveat must be heeded though: if we are to seek design targets for future mashups tools, we should be wary of looking at the artifacts resulting from existing tools as they will only show us the boundaries of what we can currently do. Looking at mashups constructed by programmers takes us past existing tool limitations but we must also be aware that programmers may not necessarily have the same needs as end-users. The more elusive targets are those applications that end-users have never constructed and beyond, to applications that end-users didn't even think were possible.

5. REFERENCES

- [1] Gamma, E., R. Helm, R. Johnson, J. Vlissades. 1995. Design patterns elements of reusable object-oriented software. Addison-Wesley professional computing series. Reading, Mass: Addison-Wesley.
- [2] Hartmann, B., L. Wu, K. Collins, and S. R. Klemmer. 2007. Programming by a sample: rapidly creating web applications with d.mix. In Proc. of UIST '07, 241-250. ACM.
- [3] Van Duyne, D.K., J.A.Landay, , and J.I. Hong, 2003. The design of sites patterns, principles, and processes for crafting a customer-centered Web experience. Boston: Addison-Wesley.
- [4] IBM QEDWiki. <http://services.alphaworks.ibm.com/qedwiki/>
- [5] Microsoft Popfly. <http://www.popfly.ms>
- [6] Nardi, B. A. 1993. A Small Matter of Programming. Cambridge: MIT Press.
- [7] Tuchinda, R., P. Szekely, and C. A. Knoblock. 2008. Building Mashups by Example. In Proceedings of IUI 2008: ACM Press.
- [8] Wong, Y., and I. Hong. 2007. Marmite: Towards End-User Programming for the Web. In Proc. of CHI '07: ACM Press.
- [9] Yahoo Pipes. <http://pipes.yahoo.com>