Exploring Websites through Contextual Facets

Yevgeniy Medynskiy
GVU Center
Georgia Institute of Technology
eugenem@gatech.edu

Mira Dontcheva
Advanced Technology Labs
Adobe Systems
mirad@adobe.com

Steven M. Drucker
Microsoft LiveLabs
Microsoft Research
sdrucker@microsoft.com

ABSTRACT

We present *contextual facets*, a novel user interface technique for navigating websites that publish large collections of semi-structured data. Contextual facets extend traditional faceted navigation techniques by transforming webpage elements into user interface components for filtering and retrieving related webpages. To investigate users' reactions to contextual facets, we built FacetPatch, a web browser that automatically generates contextual facet interfaces. As the user browses the web, FacetPatch automatically extracts semi-structured data from collections of webpages and overlays contextual facets on top of the current page. Participants in an exploratory user evaluation of FacetPatch were enthusiastic about contextual facets and often preferred them to an existing, familiar faceted navigation interface. We discuss how we improved the design of contextual facets and FacetPatch based on the results of this study.

Author Keywords
Faceted navigation, contextual facets, decision-making

ACM Classification Keywords
H.5.2. User Interfaces: Interaction styles; H.3.3 Information Search and Retrieval

INTRODUCTION

The World Wide Web is transforming into a complex environment of rich internet applications that allow individuals to browse and buy products, share reviews and recommendations, and explore and engage with new social experiences. Advances in search and navigation interfaces provide users with sophisticated tools for exploring websites and finding content of interest. Many users are also constantly appropriating these tools to suit their personal needs and habits [15,16]. To introduce the domain of this paper, we begin with a routine web browsing scenario, as it may unfold on the web today.

Sally is looking to buy a new digital camera and starts her search with her favorite online retailer. Since she is not quite sure what kind of camera she wants to buy, Sally first clicks on the "Digital Cameras" category on the navigation bar, which brings her to a page with a wide selection of cameras and some options for further narrowing the list of options. She clicks on the price range that fits her budget and begins scanning the list of cameras. The list shows the brand and model number, a small picture, and the retail price for each camera. Sally clicks through to the detailed descriptions of cameras that catch her eye. She takes note of the differences in specifications and features and other customers' reviews. None of the cameras at the top of the results list look compelling to her, and Sally decides to look at some of the more expensive but probably better cameras. She goes back to the list of available options, selects the compact form-factor category, and looks at some more cameras in detail. On one of the detail pages she notices a recommendation for a more expensive digital SLR from the same brand. Sally remembers that her brother has wanted an SLR for a long time. She decides to learn more about this type of camera and follows the hyperlink.

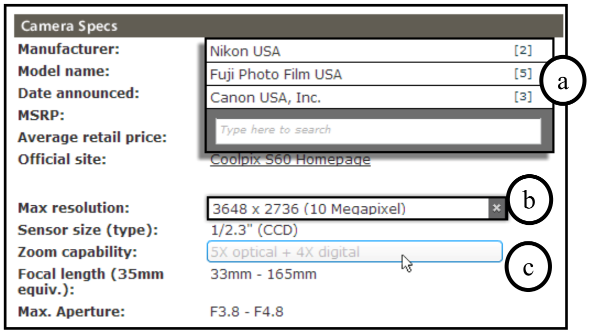
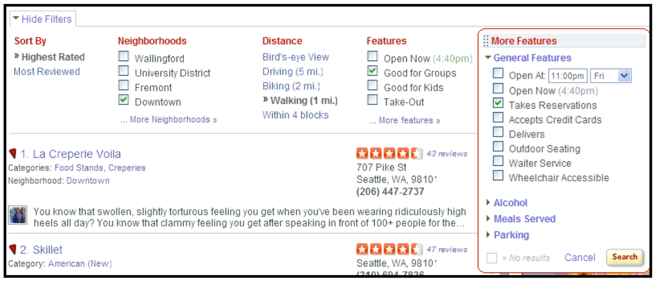

Camera Specs
Manufacturer:
Model name:
Date announced:
MSRP:
Average retail price:
Official site: Coolpix S60 Homepage
Nikon USA [2]
Fuji Photo Film USA [5]
Canon USA, Inc. [3]
Type here to search
Max resolution: 3648 x 2736 (10 Megapixel) ×
Sensor size (type): 1/2.3" (CCD)
Zoom capability: 5X optical + 4X digital
Focal length (35mm equiv.): 33mm - 165mm
Max. Aperture: F3.8 - F4.8

Figure 1. Overview of a contextual facet interface. (a) Drop-down list of facet values. The number of matches for each value appears in brackets. (b) Locked facet, showing selected value and the "remove" button. (c) Available facets in a webpage are highlighted under the cursor.

Sally's experience is quite common. When making decisions online about what to buy, where to eat, or how to spend their time, individuals rarely start out with all of their requirements spelled out ahead of time. Thus, one typically goes back and forth between a search results page and the result items' detailed description pages. As the requirements change, users may expand or narrow their search criteria and revisit items seen previously to re-evaluate them in a new light [16].

Faceted navigation [6,19] is a popular interaction technique for searching and browsing large collections of related

CHI 2009, April 4–9, 2009, Boston, MA, USA.
Copyright 2009 ACM 978-1-60558-246-7/08/04…$5.00



▾ Hide Filters

Sort By
» Highest Rated
Most Reviewed

Neighborhoods
□ Wallingford
□ University District
□ Fremont
☑ Downtown
... More Neighborhoods »

Distance
Bird's-eye View
Driving (5 mi.)
Biking (2 mi.)
» Walking (1 mi.)
Within 4 blocks

Features
□ Open Now (4:40pm)
☑ Good for Groups
□ Good for Kids
□ Take-Out
... More features »

More Features
▾ General Features
□ Open At: 11:00pm Fri
□ Open Now (4:40pm)
☑ Takes Reservations
□ Accepts Credit Cards
□ Delivers
□ Outdoor Seating
□ Waiter Service
□ Wheelchair Accessible
▸ Alcohol
▸ Meals Served
▸ Parking
□ = No results Cancel Search

1. La Creperie Voila
Categories: Food Stands, Creperies
Neighborhood: Downtown
43 reviews
707 Pike St
Seattle, WA, 98101
(206) 447-2737

You know that swollen, slightly torturous feeling you get when you've been wearing ridiculously high heels all day? You know that clammy feeling you get after speaking in front of 100+ people for the...

2. Skillet
Category: American (New)
47 reviews
Seattle, WA, 98101

Figure 2. The native faceted navigation interface for Yelp. The category filters in the "More Features" pane are initially hidden.

items. Faceted navigation interfaces expose multiple orthogonal categories (facets) of data or metadata describing a collection of items, and allow users to both refine and expand a query along these categories (Figure 2). In the scenario above, Sally browsed a collection of digital cameras using a faceted navigation interface that exposed categories such as brand, price, and form-factor. Faceted navigation interfaces are nearly ubiquitous on retail websites like Amazon.com, image repositories like GettyImages[1], and even dating websites like Match.com. Companies such as Endeca[2] and Mercado[3] specialize in building custom faceted navigation websites for businesses.

In this paper, we propose *contextual facets*, a novel interaction technique for faceted browsing. Contextual facets transform static webpage elements into user interface components that can be used to filter and retrieve similar webpages, without using a separate query interface (Figure 1). This technique offers users several key advantages over existing faceted navigation interfaces:

- The navigation UI is embedded into item detail pages, which eliminates the need to learn a separate interface for searching or browsing items.
- Users can immediately browse or filter by personally relevant categories as they notice them on item detail pages, instead of potentially having to look for less popular categories in hidden search page elements (Figure 2) or "Advanced Search" interfaces.
- They provide an effective method for navigating a website for users who land directly on an item detail page (e.g., by following a "deep link").

By turning static webpage elements into navigation components, contextual facets leverage the existing, familiar layout of item detail pages to provide users with a simple interface for accessing a broad range of metadata categories. They also provide a method for navigating websites by starting at a site's detail pages, which is helpful for users who access an item detail page directly (e.g., from the results page of a third-party search engine). Traditional faceted navigation interfaces are usually designed for users who begin their exploration from the website's front page [6,18,19], and thus are less helpful to users who arrive at the website through search engines. With contextual facets, however, users can immediately begin navigating to other alternatives.

To understand how contextual facets might be used on existing websites, we implemented FacetPatch, a prototype web browser that can automatically extract faceted data from collections of webpages and generate contextual facet interfaces. We evaluated FacetPatch with 12 subjects in an exploratory user study and found that many of them preferred contextual facets to a traditional faceted navigation interface. After presenting our results, we will discuss how we incorporated user feedback into an improved design of contextual facets and more powerful capabilities for FacetPatch. In particular, we describe how contextual facets can enable *lateral browsing*—quickly moving between similar item detail pages by manipulating item attributes.

The contributions of this work include a novel search and navigation interface through contextual facets, an algorithm for automatically extracting faceted data from collections of webpages, and design recommendations for incorporating contextual facets into web browsers and websites.

In the next two sections we describe faceted navigation and contextual facets in more detail. We then introduce FacetPatch, present the results of our user study, and describe how we incorporated these results into our designs. Finally, we present details of FacetPatch's implementation.

FACETED NAVIGATION & RELATED WORK

Faceted navigation interfaces have proven to be successful at supporting both exploratory and directed search in a variety of domains [6,10,17,18] and are nearly ubiquitous on popular web destinations. These interfaces allow users to navigate collections of items along multiple descriptive categories. In the case of musical albums these categories might be musical genre, artist, and year of release; for academic works they may include the publication type, year, venue, and author.

Hearst *et al.*'s Flamenco system [6] set the general interaction pattern for faceted navigation interfaces. Users construct search queries by selecting desired values for one or more categories, which return a result list of items matching the criteria. Clicking on an item in the result set takes the user to a detailed description page, which lists all of the available information about the item. From the detail page, users can return back to the results list or begin a new query by clicking on one of the item's attributes (Figure 3, left). For example, book details pages on booksellers' websites often allow users to click an author's name to begin a search for other books written by that author.

[1] http://www.gettyimages.com

[2] http://www.endeca.com

[3] http://www.mercado.com

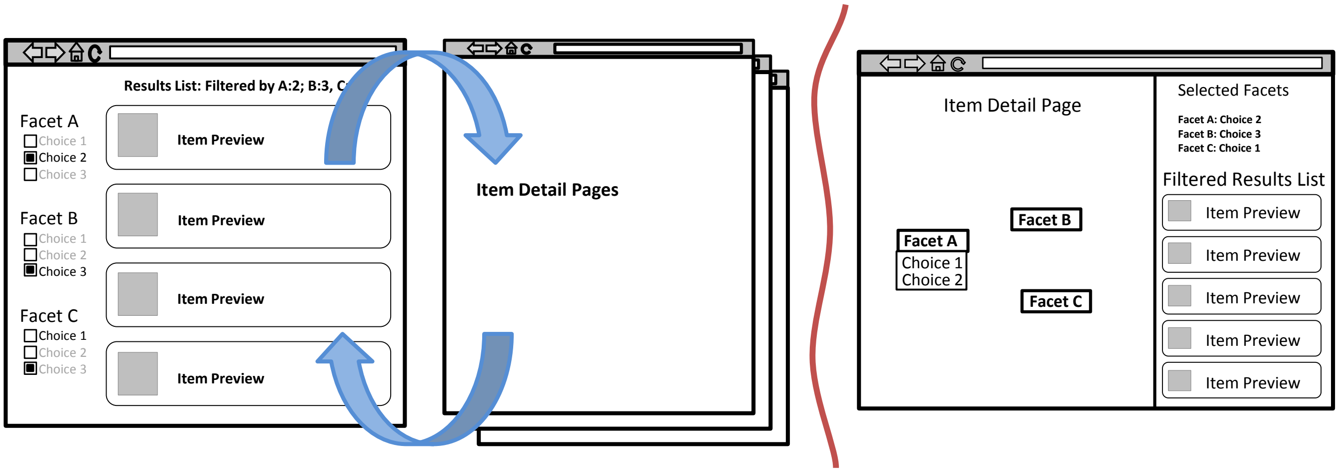


Figure 3. *Left:* Traditional faceted navigation system. The user alternates between manipulating a result set and viewing details for specific items. *Right:* Faceted navigation system with contextual facets. The user manipulates contextual facets in place on an item details page.

Faceted navigation interfaces that follow this traditional interaction pattern suffer from several drawbacks. First, the separation between constructing queries and viewing the details of any particular item can make it difficult to determine which item attributes can be used as part of a search. Second, since an exhaustive list of all available categories would make a search interface unwieldy, designers often choose to put only the most generally relevant categories directly into the interface and hide the rest in a separate interface that must be explicitly opened. Finally, a more subtle problem inherent in traditional faceted navigation interfaces is that they presume that users reach item detail pages by starting at the website homepage. However, users may land on an item detail page directly, by following “deep links” from emails or blog posts. Many users also arrive at item detail pages directly from search engines, comparison shopping websites, or other aggregators. If users bypass a website’s navigation system, they arrive at what Yee *et al*. term the “endgame” of the interface, without passing through its “opening” or “middle game” [19]. This leaves users without a clear view of how to retrieve other similar items from the same website.

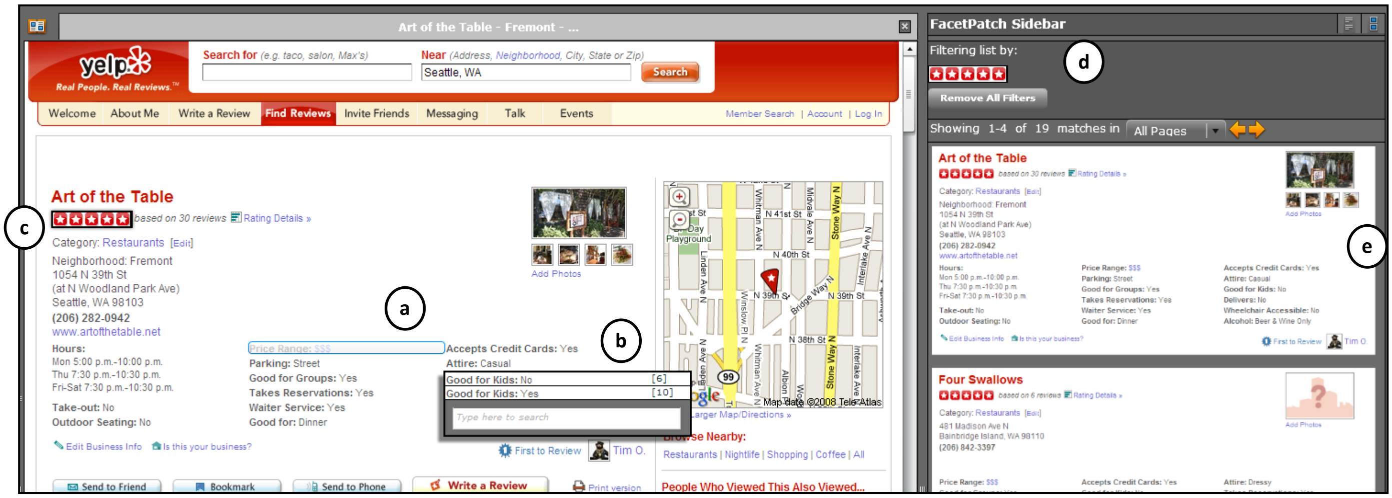
Several improvements to the traditional faceted navigation interface are described in related work. Wilson, André, and schraefel [17] introduce backwards highlighting, which helps users keep track of associations between facets in a directional faceted browser. Huynh, Miller and Karger have developed Exhibit [8] and Sifter [9], two systems that automatically generate custom faceted navigation interfaces for browsing semi-structured data. Exhibit creates a sophisticated faceted navigation interface for preexisting datasets in JSON format. Sifter is a browser plug-in that creates custom faceted filtering interfaces for browsing search results, like those returned by shopping websites. Our work differs from this previous work in two important ways. First, we provide a full faceted navigation interface on the detail webpages rather than through a separate search result webpage. Second, our query-building interface is displayed in place, on top of the existing page, rather than in a sidebar. FacetPatch and Sifter are in many ways complementary and could be combined into one navigation system for browsing both detail and search result webpages.

Many other research systems provide innovative interactions with semi-structured data in webpages. Zoetrope [1] allows users to visualize and interact with webpages as they change over time through lenses that are similar to our contextual facets. There are also a variety of mashup systems, such as MashMaker [4], that provide tools to describe semi-structured data in webpages and remix it through web services.

CONTEXTUAL FACETS

We extend faceted navigation interfaces through contextual facets, which expose faceted navigation in the context of item details pages (Figure 3, right). A *contextual facet* is an interface component that allows a user to view and select alternative values for a particular part of a webpage. Selecting a value for a contextual facet initiates a query for webpages that contain the selected value. For example, with contextual facets, a user viewing the details of a particular digital camera can click the “Manufacturer” attribute to see other manufacturers whose cameras are available on the site, click the camera’s megapixel specification to see the megapixels of other available cameras, or click the camera’s stock image to look through pictures of other available cameras (Figure 1).

The advantages of contextual facets can be described in terms of the theoretical framework for direct manipulation proposed by Hutchins, Hollans, and Norman [7]. By merging the query-building interface into the item details interface, contextual facets reduce the cognitive load required of the user to translate his or her goals into input for the faceted navigation interface, which helps the user bridge the *gulf of execution*. Users can also quickly evaluate whether an attribute displayed on an item detail page can be incorporated into a query, and then immediately use it to that end. In the terminology of Hutchins, Hollans, and Norman [7], this decreases *semantic distance* by helping the user establish whether an interface element will help him or her reach a desired goal, and decreases *articulatory*



Near: Seattle, WA

Art of the Table
★★★★★ based on 30 reviews Rating Details »
Category: Restaurants [Edit]
Neighborhood: Fremont
1054 N 39th St
(at N Woodland Park Ave)
Seattle, WA 98103
(206) 282-0942
www.artofthetable.net

Hours:
Mon 5:00 p.m.-10:00 p.m.
Thu 7:30 p.m.-10:30 p.m.
Fri-Sat 7:30 p.m.-10:30 p.m.

Take-out: No
Outdoor Seating: No

Price Range: $$$
Parking: Street
Good for Groups: Yes
Takes Reservations: Yes
Waiter Service: Yes
Good for: Dinner

Accepts Credit Cards: Yes
Attire: Casual

Good for Kids: No	[6]
Good for Kids: Yes	[10]

Type here to search

Filtering list by: ★★★★★

Showing 1-4 of 19 matches in All Pages

Art of the Table
★★★★★ based on 30 reviews Rating Details »
Category: Restaurants [Edit]
Neighborhood: Fremont
1054 N 39th St
(at N Woodland Park Ave)
Seattle, WA 98103
(206) 282-0942
www.artofthetable.net
Hours:
Mon 5:00 p.m.-10:00 p.m.
Thu 7:30 p.m.-10:30 p.m.
Fri-Sat 7:30 p.m.-10:30 p.m.
Take-out: No
Outdoor Seating: No
Price Range: $$$
Parking: Street
Good for Groups: Yes
Takes Reservations: Yes
Waiter Service: Yes
Good for: Dinner
Accepts Credit Cards: Yes
Attire: Casual
Good for Kids: No
Delivers: No
Wheelchair Accessible: No
Alcohol: Beer & Wine Only

Four Swallows
★★★★★ based on 6 reviews Rating Details »
Category: Restaurants [Edit]
481 Madison Ave N
Bainbridge Island, WA 98110
(206) 842-3397
Price Range: $$$
Accepts Credit Cards: Yes
Attire: Dressy

Figure 4. FacetPatch interface with contextual facets. (a) Highlighted contextual facet. (b) Drop-down list showing values for a facet. (c) Locked facet. (d) Additional view of currently locked facets. (e) Portion of the results list, showing thumbnail previews.

distance by minimizing the required effort to express this goal to the system.

THE FACETPATCH BROWSER

To explore how contextual facets might be used, we developed FacetPatch, a prototype web browser that can automatically generate faceted navigation interfaces with contextual facets for a collection of webpages. FacetPatch automatically recognizes webpage elements that may be categorical data or metadata and transforms them into contextual facets. FacetPatch can generate facets in two different ways. First, it can preprocess a collection of webpages for a given website and possibly store the potential facets on a server, thereby providing contextual facets to anyone who visits the website. Alternatively, it can be deployed as a browser extension and process the collection of webpages incrementally and on the fly using a user's browsing history, thereby generating facets for a user's private browsing history.

Because FacetPatch extracts faceted data automatically, without information about the semantic structure of webpages, it treats a wide range of page elements as facets. This allows FacetPatch users to navigate webpages by attributes that may be relevant in a particular context, but which are not incorporated into the navigation system provided by the website. FacetPatch can thus also automatically create a faceted navigation interface for websites that do not have a native one.

Interface

FacetPatch functions like a general purpose web browser, with the addition of a sidebar for viewing results from a contextual facet query (Figure 4). The sidebar activates anytime the user navigates to a webpage FacetPatch recognizes as an item details page. When active, the sidebar is populated with a list of other similar item details pages. For example, browsing to an Internet Movie Database (IMDb)[4] movie page populates the sidebar with a list of other IMDb movie pages, and browsing to a recipe on allrecipes[5] populates the sidebar with other allrecipes recipe pages. The user can choose to limit these pages to only those in his or her browsing history by selecting this preference from a drop-down in FacetPatch's sidebar.

Each item in the sidebar is displayed as a webpage preview thumbnail (Figure 4e). The preview thumbnails were inspired by the card metaphor introduced by Dontcheva *et al*. [2] and are an alternative to the kinds of item summaries that are usually provided by traditional faceted navigation interfaces. The user can also opt to view the list of items in the sidebar using webpage titles, as is common in traditional bookmark and history views. Let's follow Sally as she uses FacetPatch to find a restaurant.

Sally starts her restaurant exploration at a local business review website. When she arrives at a detail webpage, the FacetPatch sidebar becomes active and the webpage becomes sensitive to cursor motion. As Sally moves the mouse, any contextual facet underneath the cursor is highlighted (Figure 4a). She can click on a contextual facet to open a drop-down menu that shows the available choices. For example, in Figure 4b Sally has clicked on the contextual facet for the "Good for Kids" attribute of the restaurant and can see that there are six restaurants available that are not good for kids and ten that are good for kids. When Sally selects one of the options, her choice locks that contextual facet, draws a black outline it (Figure 4c), and filters the list in the sidebar to match the selected criteria. Sally can click on the restaurant thumbnails in the sidebar (Figure 4e) to navigate directly to other restaurant detail pages. As Sally looks at different restaurants, locked facets are animated to their new positions on the page.

[4] http://www.imdb.com

[5] http://www.allrecipes.com

Facet positions may be different on different webpages due to structural differences between pages. Sally can also see the facets that she has locked at the top of the sidebar (Figure 4d). She can remove locked facets from the search criteria by clicking the "remove" button on a particular locked facet or the "Remove all filters" button. With FacetPatch, Sally never has to leave the familiar environment of the restaurant details pages and can browse by any categorical data that is relevant to her.

EVALUATING CONTEXTUAL FACETS

To evaluate the potential of contextual facets as an interface for faceted navigation, we carried out an exploratory qualitative user study. Because faceted navigation systems are ubiquitous, we expected that our participants would already have strategies for accomplishing tasks with traditional faceted navigation systems. Thus, we were interested in observing users' existing strategies with a familiar task on a familiar website and understanding how users fit contextual facets into their framework of existing strategies. The specific questions we posed for the evaluation were:

- Can users navigate with contextual facets? Do they understand the relationship between the main FacetPatch browser and the results sidebar?
- Do contextual facets address the shortcomings of traditional faceted navigation? Do they help users adapt to changing requirements more easily?
- How do contextual facets interact with users' existing expectations for the capabilities of web browsers, such as tabs and web history?

Participants

We conducted the evaluation with a total of 12 participants. Five of the participants were female (seven male). Their ages ranged from 24 to 51, with a median of 35. Four of the participants were recruited through an internal Adobe distribution list. The remaining eight were recruited externally through an online advertisement. The participants represented a range of both technical and non-technical occupations, including a university student, a writer, and a paralegal professional. None of the participants had ever seen contextual facets or the FacetPatch system before. All of the participants were compensated with a $25 gift card.

Methodology

We chose to compare our contextual facet interface to the native faceted navigation interface of Yelp[6], a major business review website in the United States. The site is very popular in the Bay Area where we conducted the study, which made it easy to find participants already familiar with its functionality and interface. Also, the collection of webpages that make up the site—information and reviews of local restaurants, bars and other businesses—provided a rich set of faceted data that we could use to construct engaging and open-ended tasks.

We used a screener questionnaire to ask potential participants about their use of Yelp and their familiarity with various US cities. All of our participants had used Yelp previously. Ten reported visiting Yelp at least once a month ("frequent users"), and two participants reported having visited it only once or twice ("infrequent users").

We used a within-subjects experimental design. Each participant was asked to complete two tasks: one with a version of the FacetPatch in which contextual facets and the results sidebar were not available ("control condition"), and one in which they were available ("experimental condition"). We independently varied the order in which the contextual facet interface was introduced and the order of the tasks. We designed two similar tasks that served as approximations of real decision-making scenarios. In one task, the participants were asked to help their boss, who is traveling and wants to take clients out to dinner. The criteria changed throughout the task, as the boss found out more about her clients. Participants had to make sure their previous recommendations were still appropriate, and come up with alternatives for where the party could dine. The second task was very similar in structure but was about helping conference organizers decide on a list of recommendations for local restaurants and bars to be made available to conference attendees. Again, the criteria changed several times and the participants had to re-evaluate and revise their lists of recommendations. To ensure that all participants started with an equal level of knowledge about local businesses, we designed the tasks for Seattle, WA, a city with which all of our participants reported little or no familiarity.

Participants completed the tasks in one-hour sessions, with one researcher leading the session and another researcher taking notes in an observation room. The participants were told that the tasks were meant to be open-ended and that they were free to incorporate their own ideas and judgment into the solutions. The researcher mediating the session asked participants to follow a "think-aloud" protocol—describing their in-the-moment thoughts, questions, and goals as they completed the tasks.

At the beginning of each session, the participant filled out a short questionnaire about his or her general web browsing habits. Each task lasted approximately half an hour, and the participant completed a short questionnaire after each one. Participants started each task on the Yelp's Seattle homepage, with an empty browsing history. The researcher mediating the session introduced contextual facets and the features of FacetPatch before beginning the experimental condition. At the end of the session, the researcher discussed the participants' experiences with them and asked for their thoughts on FacetPatch and contextual facets.

In preparation for the user study, we made several changes to FacetPatch that facilitated the evaluation. In order to

[6] http://www.yelp.com

speed up performance, so that there wasn't any lag when a webpage was loaded, we limited the facets FacetPatch extracted to include only facets appearing in the top 400 pixels of the webpage, which is where the main restaurant attributes (such as title, average rating, photo, category, address, etc.) appear. We also pre-populated FacetPatch with the details pages of the 500 most-reviewed Seattle restaurants and bars. This pre-processed set allowed participants to navigate a wide range of establishments, while keeping system latency to a minimum. During the study, if a participant navigated to a business review page that was not in the collection, FacetPatch would process the page on the fly and add it to the collection. We expect that optimizing the extraction algorithm will improve the performance of extracting all facets, and the system will be able to handle much larger collections and sets of facets without any lag.

In the next section we describe the results of this study. We then discuss how we modified contextual facets and the FacetPatch system in response to the findings.

FINDINGS

By observing the participants' search and navigation strategies in the control condition, we were able to learn more about existing practices during decision-making tasks. The control condition also allowed us to qualitatively compare the participants' performance with Yelp's native faceted navigation interface to their performance with our contextual facet interface.

Results overview

The participants found both tasks to be of equal difficulty, and they reported being equally satisfied with their solutions to both tasks. After completing the experimental condition, they reported that using contextual facets moderately helped them to make better decisions, more quickly. However, all of the participants experienced at least some confusion with contextual facets in FacetPatch. At the same time, most of the participants, including frequent users, also had difficulties with Yelp's native faceted navigation interface.

Most of the participants (11 of the 12) were able to successfully use contextual facets to filter the list of results in the sidebar, and several participants used it as their primary or only interface for navigating Yelp in the experimental condition. Two thirds of the participants expressed at least some enthusiasm for contextual facets and FacetPatch. Some of the reasons the participants liked the interface include:

- It provided a better interface to browsing and searching Yelp. Participants who cited this reason were either not familiar with Yelp's native faceted navigation interface or had trouble finding particular filters.
- It allowed them to search directly from the details page of a particular establishment, without going back and forth between detail pages and the search results page.
- It led them to explore many different options faster.
- It included nicely formatted and informative thumbnail previews in the sidebar.

However, the other third of the participants saw no additional value to FacetPatch or to navigating with contextual facets. To them, contextual facets were a competing interface to the functionality which already exists in the native Yelp interface. It is understandable that these participants preferred an interface they already knew well. In our Discussion section, we explore how contextual facets can bring additional functionality to websites with existing faceted navigation interfaces.

Results for the control condition

During the study, most participants had at least some problems with Yelp's native faceted navigation interface. Four of the participants did not initially realize that they could filter search results using the native faceted filtering interface, even when it was expanded at the top of the search results page. Instead, they tried appending criteria, such as "takes reservations," into the text search box to narrow down their search results. Additionally, all but two of the participants had trouble finding the filters for categories like "Takes Reservations," "Live Music" or "Happy Hour", which need to be accessed by clicking a "More Features" link in the filtering interface (Figure 2). These participants mentioned that they had never used this part of the Yelp interface before, and some were surprised at the number of additional filters made available there. Most of the ten participants who used Yelp frequently mentioned that they do not typically narrow their results, except by location/neighborhood or restaurant type.

These findings provide further support for the drawbacks of faceted navigation interfaces we discussed earlier in the paper. Namely, users frequently overlook some of the functionality of faceted query-building interfaces, especially if the designers hide a subset of the available filters. Also, some users do not use a website's navigation interface. Instead, they arrive directly at the content pages by following links from outside sources, e.g., search engines. Two of the participants who used Yelp frequently said that they do not usually use Yelp's searching and browsing interface. Instead, they use Google to search Yelp, either by including "yelp" in their query string, or by opportunistically clicking links to the site when it appears in the search results.

Results for the experimental condition

Most participants started the task in the experimental condition by using Yelp's native faceted navigation interface, and only later tried contextual facets. Half of the participants started using the contextual facet interface when they became dissatisfied with a set of results from Yelp's faceted navigation interface. All participants were successful in using contextual facets, and many were eventually enthusiastic about the new interface. One user mentioned that he liked modifying parameters "on the fly."

Another said that it made her search "more organized," because she did not have to hunt around in the native filtering interface. Many felt that contextual facets allowed them to explore a variety of restaurant options faster.

All participants preferred the FacetPatch preview thumbnails over the search results on Yelp's native results page. Many participants also switched from using multiple browser tabs, a common practice to look over multiple options in the control task, to using the sidebar with the preview thumbnails in a similar manner. Seven participants wanted to be able to manipulate the results list like a personal workspace: deleting and rearranging the thumbnails, exporting the list into an email, or saving it for future reference. Two participants additionally wanted to be able to view the results in the sidebar arranged on a map, as is possible on Yelp's native search results page.

Although the participants were generally enthusiastic, they had many suggestions for improvements. Following the study, we redesigned the FacetPatch browser to incorporate their feedback.

Discoverability. Some of the participants noted that the contextual facets were "easy to forget about" and did not stand out from the page. We modified our preliminary design, which was fairly subtle, to make contextual facets appear more like buttons (Figure 4a). In the next section, we discuss the tradeoffs around discoverability. Some participants suggested that the facets should be on the side, similar to the query-building interfaces of many websites, even though this may bring about long sidebars which may require scrolling. Several participants also suggested adding filtering by a range of values for some categories, such as a price range or average review rating.

Feedback. Since most participants started their exploration through Yelp's native interface, they were unclear about the relationship between the list of results in FacetPatch's sidebar and the Yelp search results. For example, while looking at a Thai restaurant that did not take reservations, one participant clicked the "Takes Reservations" contextual facet in the page, and selected "Takes Reservations: Yes." Looking through the results that this action returned, he was surprised that most of them were non-Thai restaurants. This participant, and most others, initially expected that FacetPatch was directly integrated with Yelp's searching and browsing interfaces. Most quickly realized that their selections did not transfer from one interface to the other, at which point they proceeded to lock additional contextual facets (such as "Category") to narrow the search to the desired granularity. To more clearly distinguish the contextual facet query as separate from any query performed in Yelp's native interface, we added a listing of the active contextual facets on top of the results sidebar (Figure 4). Thus, users are able to more easily tell that the restaurants were not filtered by, e.g., "Category: Thai."

In our preliminary design locking contextual facets did not trigger automatic navigation to new webpages. To load a new webpage that matches locked facets, the user had to click an item in the sidebar results list. Thus, it was possible to "create a view" of a webpage that was not actually accurate, which was confusing to some participants. For example, one participant locked a contextual facet to "Happy Hour: Yes" on the details page of a bar which listed "Happy Hour: No." A short while later, while still on the same page, he mistook the locked facet for the page's original content and noted that the bar offered a happy hour. In our redesign, locking a contextual facet navigates the user to a new page that fits all of their locked criteria.

State persistence. The persistence of queries was a major concern for many of the participants. In the version of FacetPatch we evaluated, if users navigated away from a business detail page, any locked contextual facets were cleared. Thus, if a participant clicked through to photos of a restaurant from its details page, which loaded pages that FacetPatch does not recognize as business listings, he or she would have to recreate any previously set contextual facet query. We resolved this in our redesign, as it was a major usability hurdle.

Thumbnail ordering. Although the participants were very enthusiastic about the thumbnails, many were confused about their ordering. Part of the confusion was due to the already mentioned inconsistency with the Yelp search interface. In our redesign of the FacetPatch browser, we added a new ranking function that sorts the results using similarity to the current webpage. Pages most similar to the currently active one are listed first in the sidebar, which allows users to retrieve restaurants that are similar to the one they are currently viewing but, better fit their criteria.

Browsing history. The integration of browsing history into the result set in the sidebar was not very popular, because it showed the participants *all* of the business details pages they had previously visited, which included both ones they liked and ones they didn't like. Additionally, the history view changed as participants tried to use it—clicking a page to revisit it caused it to move to the beginning of the list, throwing off the order of the rest of the items. In the post-evaluation discussion, two participants said that they would like to have the option of removing items they had seen before from result sets on Yelp or other websites. One participant suggested that the history view would be more useful if it showed him how he arrived at a particular page: whether he navigated to it from a list of search results, by following a hyperlink on another page, or otherwise.

DISCUSSION

The user study allowed us to see how individuals might integrate a contextual facet interface into their existing browsing strategies on a website that was familiar to them. Although they encountered a variety of problems, nearly all were able to use contextual facets to complete the task in the experimental condition. Two thirds of the participants were enthusiastic about the system, and felt it would be useful to integrate it into other websites they used frequently, such as the auction site eBay, the dating site Match.com, and shopping sites like Amazon.com. The other third of participants had more reservations about the system, because they weren't sure how it moved beyond the existing, familiar interface provided by Yelp. In this section, we discuss how we improved contextual facets to provide functionality that is not available in traditional faceted navigation interfaces.

Integrating contextual facets into the web UI

One of the challenges that surfaced in the evaluation was the discoverability of contextual facets. Contextual facets layer new functionality onto the existing "web UI"—the collective interface and functionality provided by web browsers and websites—by embedding a set of interactions into webpage elements, which may or may not have been interactive previously. Browser plug-ins that layer functionality on top of existing webpages generally work in one of two ways: heavyweight interactions are usually modal and are initiated through a toolbar (e.g., Web Summaries [3]) or a sidebar (e.g., Sifter [9]), while more lightweight interactions are initiated through interactive elements embedded into the webpage itself (e.g., Cooliris Previews[7]). Since contextual facets are fairly lightweight UI components, we embedded them into the page, and tried to make them unobtrusive.

While a browser plug-in or a separate browser provides a general solution for enabling faceted navigation through contextual facets, an alternative approach would be for a website's designer to add contextual facets to the site's interface. The designer may be in a better position to decide how the low-level mechanics of the contextual facet interface are to interact with the website's particular interface, because he or she has a better understanding of visitors' needs and behaviors.

Integrating contextual facets with web browsing

During our study, we found that participants misunderstood or overlooked visual feedback about the navigation state provided by searching and browsing interfaces relatively often. This was the case during both the experimental and control conditions. On two occasions, a participant did not notice that his or her search was narrowed down to a particular neighborhood. On Yelp, the neighborhood filter is shown in a breadcrumb trail at the top of the page (a common practice on other websites as well), but neither participant directed his or her attention there to try to diagnose why a search was returning seemingly too few results. Some participants also had an inaccurate model of how Yelp's native search filtered results. Participants who started the personal assistant task by searching for "thai" from the search box on Yelp's homepage would frequently assume that the results contained only Thai restaurants. However, the results contained all businesses with the word "thai" anywhere on their details page, and some of the participants became confused or frustrated if a combination of filters they then selected left them with results that were not Thai, not restaurants, or both. This suggests that navigation interfaces need to be more transparent on their filtering and search criteria. We improved FacetPatch by adding a special area in the interface specifically for this purpose (see Figure 4d).

The transience of selected filters is a general problem that often leads to confusion around search results. During the pre- and post-evaluation interviews, many participants described performing the same queries frequently on a range of websites. Seven of the ten frequent Yelp users agreed or agreed strongly with the statement "On Yelp, I visit the same kinds of pages more than once," which we asked in the introductory questionnaire. Thus, users would likely benefit from being able to save sets of frequently used filters on websites more generally. A system like SearchBar [13] can help users retrieve previously visited webpage not only by keyword queries but also by previously activated filters.

Integrating contextual facets with decision-making

A key trend we observed during the evaluation was that in the experimental condition most participants did not begin searching or browsing Yelp with contextual facets. Instead, they began with Yelp's native navigation interface and turned to the contextual facet interface when they had trouble getting good results or when a particular result no longer fit some new requirement. Thus, many of the detail pages participants used in order to initiate a query through contextual facets were pages they wanted to leave.

To support using contextual facets in this way, we modified FacetPatch to load a new page whenever the user selects a value through a contextual facet that is different from the value on the current page. For example, if the user is viewing a restaurant that doesn't take reservations and selects the value "Takes Reservations: Yes," FacetPatch will lock the facet and load the first restaurant listed in the results sidebar that does take reservations. As a result, the user can now use contextual facets to navigate through the space of alternatives, building up a query along the way. To facilitate this type of navigation, we implemented an algorithm to calculate similarity between pages of the same type. Pages in the results sidebar are now sorted according to their similarity to the currently active page. This makes it easy to use FacetPatch to find better alternatives to a result that is "almost right."

The lateral browsing that these two changes enable could be especially useful to users who navigate to a website like Yelp using a third-party search engine. Because these users have likely already specified some of their preferences into the search engine, they may not be willing to re-specify their preferences again with a new search interface, if the result they found isn't quite right. With the current implementation of contextual facets, however, they can directly manipulate the values of the attributes that aren't satisfactory to navigate to better-fitting pages.

[7] http://www.cooliris.com/site/firefox/

IMPLEMENTATION

FacetPatch was implemented as an Adobe AIR[8] application written in ActionScript 3. The application uses the WebKit HTML rendering engine to parse, render, and manipulate webpages. We now describe two key technologies we developed as part of FacetPatch: an algorithm for automatically extracting faceted data from collections of webpages conforming to a common template, and an algorithm for computing similarity between two webpages using the values of their facets.

Automatic extraction of faceted data

Before it can generate contextual facets to overlay on top of a webpage, FacetPatch must extract faceted data from a collection of similar webpages—for example, business listings on Yelp or recipes on allrecipes. The application can extract this data from a set of webpages in a pre-processing step or compile it one page at a time as the user browses the web. Currently, FacetPatch recognizes webpages that conform to a common template by matching their URLs against a database of known URL patterns. For example, the URLs of all Yelp business listings match the regular expression *www.yelp.com/biz/.**, the URLs of all IMDb movie entries match the regular expression *www.imdb.com/title/tt[0-9]{7}/*, and so forth.

To identify a data category and extract its possible values, FacetPatch searches for recurring webpage elements whose content varies. Each page element that appears in multiple pages is treated as a category (facet), and the values it takes on are its possible values. For example, given a collection of Yelp business listings, FacetPatch will create facets corresponding to the business name, average user rating, and various attributes such as "Takes Reservations" and "Good for Kids," because each of these is provided in a uniquely identifiable, recurring webpage element.

To match webpage elements across a collection of webpages, we use several techniques for matching nodes of the Document Object Model (DOM) tree structures which describe webpages. First, we extract a list of DOM elements that may be facets by compiling a list of terminal elements which appear to contain a single unit of information, such as a short snippet of text, a single image, or a small number of both. We then match up terminal elements across webpages by looking for elements with identical ID attributes, identical paths to the top-level document node (XPaths), and identical near-by text labels. Because page elements that ostensibly show the same facet (e.g., business name) do not appear in the exact same location of the DOM tree in all pages, we combine those three matching heuristics for greater accuracy. As nodes are matched, we associate different values with the same facet, and preserve the visual style of each facet value by copying its element's CSS style along with its content. Matched nodes are only considered a facet if they take on multiple values; matched nodes with a single value across all pages are considered page chrome (e.g., the website logo).

Previous work has described techniques for extracting user-selected webpage elements across similar pages [5,11,12], the automatic extraction of webpage templates to match semantically-similar content [14], and the automatic extraction of structured data from individual webpages [9,20]. The contribution of our algorithm lies in automatically extracting structured, faceted data from a collection of webpages that conform to a common visual template, *without* requiring users to specify which content to extract. Though the heuristics we use are relatively simple, we have found them to work well across different domains. We have successfully tested our technique on business listings from Yelp, movie entries from IMDb, recipes from allrecipes, digital camera entries from the Digital Camera Resource[9], music CD entries from allmusic[10], and real estate listings from Zillow[11].

After extracting faceted data from a page, FacetPatch generates its preview thumbnail for the results sidebar. The thumbnail is generated by cropping an image of the rendered webpage, starting a few pixels above the top-most facet, setting the default height of the thumbnail to 350 pixels, and using a width that will contain all contextual facets within those 350 pixels.

Computing the webpage similarity metric

In the current implementation of FacetPatch, items in the results sidebar are ordered by a measure of their similarity to the currently active webpage. This also gives a natural ordering to the facet values in a drop-down list for a contextual facet: values appearing in webpages that are more similar to the currently loaded one appear higher.

We compute the measure of similarity between two webpages by comparing the facets they have in common. For each such facet, we compute the similarity between the facet values. The sum of these is the measure of similarity between two webpages. We have tried using two different methods to compute the similarity between two values of a facet: simple equivalence, and an approach based on string edit distance between the values' string representations. For the equivalence method, the similarity between two webpages is simply the number of facets for which they have identical values. We find that this very simple and efficient method is surprisingly effective. In our collection of Yelp business reviews, for example, the restaurants ranked most similar to a given restaurant generally offer similar cuisine, are in the same price range, etc. We are currently exploring more sophisticated metrics of similarity found in the information retrieval literature, which may be more appropriate for different collections.

[8] http://www.adobe.com/products/air/

[9] http://www.dcresource.com

[10] http://www.allmusic.com

[11] http://www.zillow.com

CONCLUSIONS & FUTURE WORK

We have described the design, development, and evaluation of contextual facets, a novel interaction technique for exposing a faceted navigation interface through item detail webpages. Contextual facets allow users to explore a collection of webpages without deferring to a separate search and browsing interface. This navigation technique can be especially useful for users who reach item details pages through general search engines, thus bypassing the native searching and browsing interface provided by the website. The feedback from our study leads us to the conclusion that contextual facets hold promise as an interface for navigating websites. In our redesign, we improved the low-level interactivity and discoverability of contextual facets. We have also adapted contextual facets to more closely fit users' decision-making processes.

We see several promising avenues for future research. One direction is to build a toolkit for integrating contextual facets into websites. Similarly to Exhibit [8], such a toolkit would make it easy for hobbyists and non-programmers to integrate a faceted navigation interface into websites that do not yet have a sophisticated navigation interface. A toolkit aimed for designers could allow them to specify the semantic structure and relative importance of different facets in their data, thus personalizing the contextual facet interface to the needs of their visitors.

We are also interested in continuing to develop FacetPatch as a system for automatically generating contextual facet interfaces, perhaps as a browser plug-in. Given the user feedback, we hope to improve the thumbnail representation of search results to better aid users in making decisions. One approach is to improve the thumbnail representation over time as the system learns which facets are personally important to each user.

ACKNOWLEDGMENTS

Many thanks to David Salesin, Wil Li, Joel Brandt, Sharon Lin, and the anonymous reviewers for their valuable comments and feedback.

REFERENCES

1. Adar, E., Dontcheva, M., Fogarty, J., and Weld, D. Zoetrope: Interacting with the Ephemeral Web. *Proc. UIST 2008*, ACM (2008), 239-248.
2. Dontcheva, M., Drucker, S.M., Salesin, D., and Cohen, M.F. Relations, cards, and search templates: user-guided web data integration and layout. *Proc. UIST 2007*, ACM (2007), 61-70.
3. Dontcheva, M., Drucker, S.M., Wade, G., Salesin, D., and Cohen, M.F. Summarizing personal web browsing sessions. *Proc. UIST 2006*, ACM (2006), 115-124.
4. Ennals, R.J. and Garofalakis, M.N. MashMaker: mashups for the masses. *Proc. MOD 2007*, ACM (2007), 1116-1118.
5. Han, J., Han, D., Lin, C., Zeng, H., Chen, Z., and Yu, Y. Homepage live: automatic block tracing for web personalization. *Proc. WWW 2007*, ACM (2007), 1-10.
6. Hearst, M., Elliott, A., English, J., Sinha, R., Swearingen, K., and Yee, K. Finding the flow in web site search. *Comm. ACM 45*, 9 (2002), 42-49.
7. Hutchins, E.L., Hollan, J.D., and Norman, D.A. Direct Manipulation Interfaces. *Human-Computer Interaction 1*, 4 (1985), 311-338.
8. Huynh, D.F., Karger, D.R., and Miller, R.C. Exhibit: lightweight structured data publishing. *Proc. WWW 2007*, ACM (2007), 737-746.
9. Huynh, D.F., Miller, R.C., and Karger, D.R. Enabling web browsers to augment web sites' filtering and sorting functionalities. *Proc. UIST 2006*, ACM, 125-134.
10. Karlson, A.K., Robertson, G.G., Robbins, D.C., Czerwinski, M.P., and Smith, G.R. FaThumb: a facet-based interface for mobile search. *Proc. CHI 2006*, ACM (2006), 711-720.
11. Kowalkiewicz, M., Orlowska, M., Kaczmarek, T., and Abramowicz, W. Towards More Personalized Web: Extraction and Integration of Dynamic Content from the Web. *Frontiers of WWW Research and Development - APWeb 2006*, Springer Berlin (2006), 668-679.
12. Lingam, S. and Elbaum, S. Supporting end-users in the creation of dependable web clips. *Proc. WWW 2007*, ACM (2007), 953-962.
13. Morris, D., Morris, M.R., and Venolia, G. SearchBar: a search-centric web history for task resumption and information re-finding. *Proc. CHI 2008*, ACM (2008), 1207-1216.
14. Reis, D.C., Golgher, P.B., Silva, A.S., and Laender, A.F. Automatic web news extraction using tree edit distance. *Proc. WWW 2004*, ACM (2004), 502-511.
15. Weinreich, H., Obendorf, H., Herder, E., and Mayer, M. Off the beaten tracks: exploring three aspects of web navigation. *Proc. WWW 2006*, ACM (2006), 133-142.
16. White, R.W. and Drucker, S.M. Investigating behavioral variability in web search. *Proc. WWW 2007*, ACM (2007), 21-30.
17. Wilson, M.L., André, P., and schraefel, M. Backward highlighting: enhancing faceted search. *Proc. UIST 2008*, ACM (2008), 235-238.
18. Wilson, M.L. and schraefel, M. A longitudinal study of exploratory and keyword search. *Proc. ACM/IEEE-CS Joint Conf. on Digital Libraries*, ACM (2008), 52-56.
19. Yee, K., Swearingen, K., Li, K., and Hearst, M. Faceted metadata for image search and browsing. *Proc. CHI 2003*, ACM (2003), 401-408.
20. Zhai, Y. and Liu, B. Web data extraction based on partial tree alignment. *Proc. WWW 2005*, ACM, 76-85.