

Multimodal Shopping Lists

Jhilmil Jain, Riddhiman Ghosh, and Mohamed Dekhil

Hewlett-Packard Labs, 1501 Page Mill Road, Palo Alto CA 94304
{jhilmil.jain,riddhiman.ghosh,mohamed.dekhil}@hp.com

Abstract. In this paper we present a prototype for creating shopping lists using multiple input devices such as desktop, smart phones, landline or cell phones and in multimodal formats such as structured text, audio, still images, video, unstructured text and annotated media. The prototype was used by 10 participants in a two week longitudinal study. The goal was to analyze the process that users go through in order to create and manage shopping related projects. Based on these findings, we recommend desirable features for personal information management systems specifically designed for managing collaborative shopping lists.

Keywords: shopping lists, images, text, video, audio, mobile, web.

1 Introduction

We define “intent” as any stage in the shopping cycle that typically starts with a generation of an impulse or a need for a product and ends in a purchase. This process can be broken down into the following intent levels - Level 1: impulse or need; Level 2: research, advice seeking; Level 3: price comparison; and Level 4: decision to buy. Please note that these levels are not linear, for example level 2 and 3 can occur in

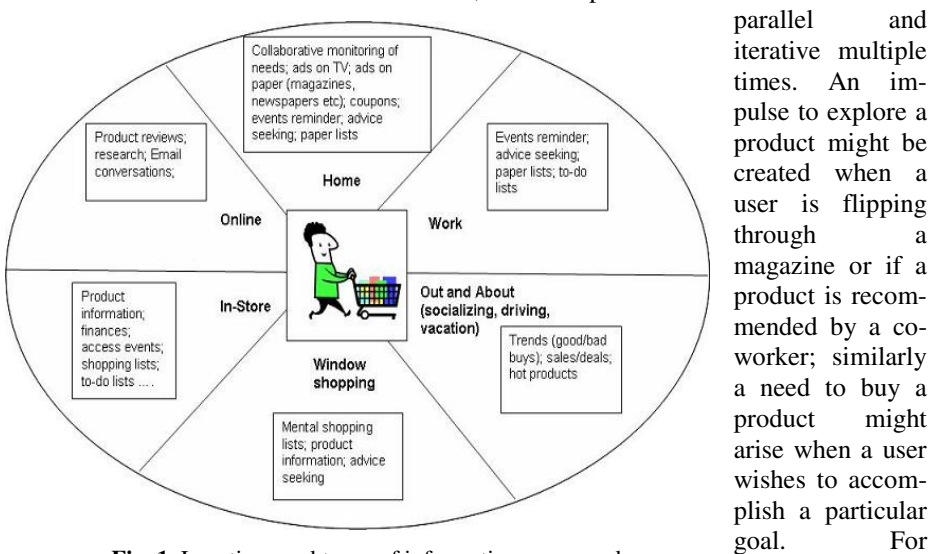


Fig. 1. Locations and types of information consumed

example, Jane is a teenager who bought her first car and noticed that she would save a lot of time while driving if she had a GPS system. Jane now enters the first level of the intent cycle. She asks her friends and family for recommendations and also searches online for a GPS that is within her price range, or visit various retail stores to test the look and feel of the product or to get expert advice; these activities shift her into level 2. This is followed by price comparisons online to determine a trustworthy store which has the best price, shifting her into level 3. Levels 2 and 3 are performed iteratively and are usually intertwined until level 4 (i.e. decision to buy a particular brand/make of a GPS) is reached.

As the above example illustrates, the task of shopping is: a) complex, requiring multiple decisions b) collaborative, involving advice and permission seeking, c) social, involving experiential and recreational aspects [6], [9]. Shopping tasks span various locations (e.g. home, store, work etc.), involve multiple touch points (e.g. online, discussions with friends and family, mobile etc.), and tools (e.g. to-do lists, email, SMS, social networking websites, reminders etc.). Figure 1 illustrates the type of information and locations of these intents that the user would like to capture Systems that require consumers to track and report their shopping intent in a structured manner are not very successful [10]. This motivates the use of systems that will allow for unobtrusive or casual capture of shopping related intents using suitable modalities.

The main issues that we want to address in this space are: a) capturing intent in multiple places/locations; b) using existing technology and devices with minimal user effort to capture and access information, c) managing different levels of intent; and d) converting intention to behavior or action. There are three types of intention-action gaps [2]: physical - not being in the right place (how can the consumer record an intent if no access to their desktop is available or if they need to capture an impulse which will be acted upon once they are in the store); awareness - not the current focus of attention (how can intent be recorded if the consumer is driving or is in a meeting); and informational - insufficient information (how can advice from friends and family be collected to make an informed decision).

2 User Survey

Our goal was to first understand the common tools used by people to organize their shopping trips, the frustrations that they face in retail stores, and missing features that they would like retailers to address. We conducted a Web-based survey and multiple follow-up interviews to identify the severity and frequency of shopping frustrations and the services that would be useful in a retail environment.

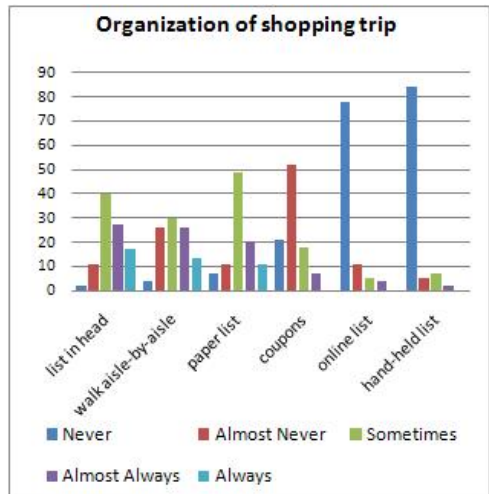


Fig. 2. Tools used to organize shopping trips

There were 69 respondents, of whom 54% were female and 46% were male. 70% were primarily responsible for shopping, and 32% had children in their household. We found that shopping trips for most consumers are usually unorganized, especially for commodity items. Figure 2 details the popularity of various tools used for organizing shopping trips by the participants in our survey.

It was observed that most participants (44%) usually make a mental note of things that need to be purchased (27% answered almost always and 17% answered always) or they walk down aisle-by-aisle in the retail store to determine items to be bought (26% answered almost always, 13% answered always).

These results are in line with Thomas's study on list making behaviors of grocery shoppers [8]. Our study also indicated (as expected) that since most shoppers are unorganized, forgetting what they came to buy was rated as one of the top 5 shopping frustrations.

3 Our Solution

We implemented a solution that enables the creation of shopping lists using different modalities or formats such as structured text, audio, still images, video, unstructured scribbles and annotated media (i.e. annotating images, video frames etc.). The lists can be created from three primary sources: 1) using a toolbar on the Web when users have access to a computer, (fig. 3 and 4) or 2) using any phone (landline or cell phone) by dialing into a number and leaving an audio message when users do not have access to the Web on the phone, or 3) by downloading a mobile client on their PDAs and accessing the service using the data network provided by the cellular carrier or any wireless network (fig. 5). These multimodal shopping lists are made accessible via the Web on any desktop (fig. 6) or mobile phone (fig. 7).

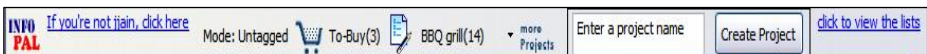


Fig. 3. Closeup of the browser toolbar

3.1 Interaction Details

Users can install a Web browser toolbar (fig. 3). During this process, the user is required to choose a username and password to use the service from a Web browser, and a phone number to use the service from a landline or cell phone. To add an item to the shopping list from the Web, the user needs to highlight the product details (this could be in the form of text, image, audio or video clip) on the Web page (fig. 4) and click on the relevant list name (e.g. "To Buy"). Behind the scenes, a request is sent to the Web server with the Web page URL and the date and time of the capture. There are no restrictions on the type of information (i.e. text/image/audio/video) or the website – thus any information from any website can be added to the shopping list. Users can collaboratively add to the same shopping list on the Web by using the same login information. Similarly, other intents such as researching an item, permission/advice seeking, reminder to buy item for a particular events, etc. can be added to the respective lists named – Research, Ask, and Events.

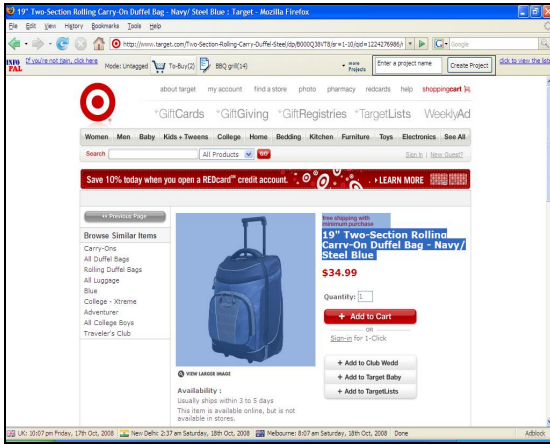


Fig. 4. To capture information (such as text and images) from the web highlight relevant information and click on the list name to add it to the list. Here we are adding to the “To-Buy” list.

Fig. 5. Mobile client to capture shopping information from handhelds

Audio input from a phone can be added to the list by calling a predefined phone-number to record a voice message. Our system extracts keywords from the audio and adds it to the original recorded audio, along with the date and time information.



Fig. 6. Multimodal shopping lists accessed from the Web. Samples above contain audio, images and text.

Fig. 7. Multimodal shopping list accessed from a mobile phone

Still images, audio and video based information are added via MMS using preexisting software on the cell phone. The email send from the phone is parsed such that the multimedia file, date, time (and optionally text) information is extracted and added to the appropriate list (the list name is the subject of the email). Alternatively, users can download and install a mobile client to add multimodal information to the shopping lists (fig 5).

The user can access the aggregated multimodal lists from any Web enabled device – on the desktop (fig. 6) or on the mobile device (fig 7).

4 Longitudinal Study and Findings

Ten users were asked to install and use the system to create lists via the Web and phone while maintaining a diary to record their thoughts and observations in a week long longitudinal study. Five of the participants were married and five were single; age distribution was fairly uniform (20s – 3; 30s – 2; 40s – 3; 50s – 2.) On an average, the participants shopped for groceries 1-2 times a week, browsed or shopped online 4 times a week; and shopped in retail stores 2 times a week. Before using our system, participants were using tools like kitchen whiteboards, post-it notes, paper lists, SMS, and email to create shopping lists.

During the week long study, a total of 255 entries were made by all 10 users. 187 entries were made on the Web and 68 entries (approximately 30%) were made from a phone. At the end of the week long study we conducted one-on-one in-depth interviews with each participant that lasted between 60-90 minutes and also asked them to complete a usability questionnaire. The interviews and participant diaries were analyzed inductively using the grounded theory approach to generate design guidelines. Using this approach, we first discovered the key components, and then identified a number of categories and concepts, as well as interrelations between each of them using coding and sorting. The grounded theory analysis identified multiple key experiential categories as described below:

Highlight and add to list

This technique of adding items to a list was extremely popular. On day 2 of use, one user said: *“added items to the cart quickly, like second nature now. [I now] know to highlight more text like price, stuff that will help me with making a decision”*. Another user said: *“I thought I had to highlight the item and then drag and drop it to the list. As soon as I realized that I can just click the list name to add the item I was thrilled.”* This mechanism provides a flexible and intuitive way to select and capture information of choice. The primary purpose of the toolbar was to allow users to add items to their shopping list that they intended to purchase at the store. Interestingly the toolbar was used for various other “unintended” purposes as well.

Comparison shopping

One of the most popular uses of our toolbar was in the form of product and price comparison from various websites. Comparison websites like PriceGrabber.com, BizRate.com, NextTag.com, etc. extract product prices from pre-configured websites that cannot be customized by users. Our toolbar helped users highlight and capture product details they were considering from any website and allowed them to prune the information on demand. *“It served as my personal customizable Froogle [1]”*

Collaborative shopping

The next popular use was to create collaborative lists. At the time of the experiment, multiple users could share the login information in order to add items to the same list.

One of our participants was the primary person responsible for weekly shopping chores in a family of four. Other family members would use email for communication whenever they thought of something to be bought. She kept a centralized list that she would update whenever a new request was made. She asked her family members to use the toolbar instead for a week and said: *“It [the toolbar] really made life easier for me – saved me a lot of time and reduced confusions”*.

Another participant used the toolbar to find a birthday gift for her niece and she said: *“It would be nice to share it [shopping list] with my sister since she is also looking for gifts for my niece (her daughter). Last night I showed her some of the things I found (from memory) but it would be nice to share the list.”*

This example illustrates that sharing of lists is a very important aspect. In the next version, we plan to add features that will allow sharing the whole or parts of the list with family/friends in real time without having to share login information. We will also have the capability using which the list owner can identify the user that added a particular item.

Organizing shopping

“[I am] finding it [toolbar] an extremely useful tool for organizing my web shopping... especially the panic, last minute b-day shopping where I move from site to site at about 100 mph. Usually I just rely on memory... I can spend a lot of time trying to remember and retrace my steps”. We observed that when users started with level 1 of their intent cycle, they rarely knew what structure their activity would take. Our toolbar was intentionally designed to be flexible and light weight so that items of interest can be quickly collected in an unstructured way and organized as the list evolves. As we move forward more sophisticated features are required to manage the lists once they are collected. We will highlight these in the section on discussions.

Editing and modifying list

One of our participants, an administrative assistant, used the toolbar to track items that were left after an office event. Initially she created a list of the items that she bought for the event, and at the end of the event she edited the list to remove the items that were consumed. She said: *“After an event I have no way of tracking and communicating [to other admins] that I have 5 cases of soda remaining that can be used for future events”*. This example illustrates the need to have functionality for editing and modifying lists.

Repeatable lists

“We use the same things for various events”. One of the participants mentioned that products used for similar events remain unchanged; yet for every event they had to make a new list. Using this toolbar she created a “master list” that was printed and shared with other admins. When they are planning a new event for the first time, they are able to simplify the process by using items from the master list to initiate a new list. This example demonstrates the need of creating repeatable lists or templates for events. These templates serve as a starting point and provide a structure to the activity being planned.

Re-finding information

The toolbar allows users to save items they are currently interested in or are currently tracking. The toolbar was used to re-find information that users had looked at a while ago. After using the tool for 4 days one user said: *“I am at the point where I am ready to make a decision about what to buy. I really like having this list. I had totally forgotten some of the things I had selected earlier in the week! I know I would not have written this stuff down, I may have bookmarked some but I am hesitant to bookmark short-term stuff because it clutters bookmarks and requires cleaning up later (which I often don't do).”*

The organization of the bookmarks is different from the information collected by the toolbar. Users spend significant amount of effort to create and maintain their bookmarks, where as there seems to be a low commitment to the upkeep of the information captured from the toolbar and its more “throw-away”. This is due to the fact that the toolbar serves a different purpose than bookmarks and the information captured is transient, and re-finding is more short-term. This could be the reason that prevents users from adding retail related information to the browser bookmarks.

Wish list

One of our participants used the toolbar to create a wish list of clothing by capturing text and images from online fashion magazines and forums with the primary purpose of finding affordable versions on her next shopping trip. She said: *“I have a word file in which I copy and paste the images and description of the clothes that I am interested in, and the website URL from which I copied them – this is so much work, and I always have to remember to email the file to myself so that I can access anywhere.”*

The ease of collecting text and images from the website, allowed her to form an initial estimate of clothes that she was interested in and she could later use this initial list to make additional decisions. It was important for her that the system could tolerate the collection of non exact items that she could use for interpretation and that the shopping list served more as a point of inspiration than a definitive list.

Research for consumers with special needs

A few of our participants or their family members had special needs (e.g. allergies to peanuts, vegan diet). Due to this they had to research each item before purchasing it from a store. They would typically go to pre-determined websites to conduct this research. Though they liked the ability of collecting related information about a product in a centralized location, the system could in future provide features to help them determine if a product is suitable for their needs by partially automating the process – e.g. extract ingredients information and flag the items which contain peanuts in the ingredients.

Reminder

Participants used the toolbar for creating location based “to-do” items – e.g. *“get quarters for laundry from Albertsons, mail letters”* by adding free-form notes to the Web shopping list. These reminders were not necessarily related to the products on the shopping list and provided the users with a low-effort mechanism of customizing the lists. This flexibility and free-form capture of the toolbar is important since it is difficult to predict what the lists will be used for.

Audio

The users typically used the audio interface while they were driving or when they were in a “hands-busy” and “eyes-busy” environment. 30% of the messages were sent via the phone, even though a few steps were required to login before the users could leave an audio message. We believe that this could be a much higher number if users could leave a message by one-touch dialing.

Most users had not expected the system to extract keywords from audio, and were pleasantly surprised. On an average, the system takes about 2-3 minutes to extract keywords from the audio message and populate the shopping list. Most users felt that this response time was satisfactory even though recognition/extraction was not accurate 100% of the time.

The audio part of the system was very useful for product comparison purposes, though the login and dialog management were tedious. Some users sent voice reminders of prices of products (e.g. multigrain bread at Albertsons was \$2.50 and at Trader Joe's was \$3.45) in order to compare them at a later stage.

"It would be nice to type in notes or suggestions about what I am looking for. Some businesses don't have a web site or don't list their products on their web site, but I know they are a good places to buy... like Mervyns (good HelloKitty stuff) and a Japanese store that has good Japanese character toys. The phone allows me to do this, so I will add these when I can connect". This example illustrates the need for creating free-form notes that users can use as reminders or for sharing their recommendations.

A drawback of using a third party solution for implementing audio capture was the lack of customization of instructions on how to use the system. Since we intentionally did not provide instructions on how to use the system to mimic a real life situation where users may not read manuals, two of the ten participants did not realize that they could leave a voice message with multiple product names in the same message. We believe that this issue could have been resolved if the "greeting" message of the audio piece was customized with this fact.

Most consumers typically do not use the Web to shop for commodity items. This could be the reason why the toolbar interface of the system was not used much for adding grocery items to the shopping lists (approx. 1%) and approximately 90% of items added from the phone were grocery items.

5 Related Work

There have been other efforts to assist users to capture purchase information to be used in a retail environment. The Domisilica [5] project proposes a refrigerator with a network connection and attached display to display and record shopping lists among other things. However such "computationally enhanced" devices are out of reach of the common customer in the real world. Also, we feel that refrigerators, even if they were computationally enhanced, would not be able to support ubiquitous, pervasive and timely capture of shopping intent. It differs from our system which uses commodity hardware and software.

Sellen et al. from HP Labs describe concepts for a reminder service targeted towards working parents, where reminders or notifications are delivered to parents inside a car [7]. This system differs from ours as the only inputs it affords for reminder creation are either "offline" at a computer, or by typing on a handheld.

PlaceMail [4] is a location-based reminder service. We find the following important differentiating factors between this system and ours: The system requires a GPS (which does not work indoors) enabled cell phone with a wireless "data" connection between a user cell phone and the system for data transfer. Our system works with both a regular POTS phone and any cell phone, and uses the cellular/telco carrier to interact with the system rather than a "data" connection. Also, using a cell phone on Placemail requires a user to login by typing on the cell phone before it can create reminders. Moreover, creating a reminder involves specifying time/location (translated to latitude/longitude). We feel this requires significant effort by the user and the interaction with the system is not through natural interfaces. This is in fact reflected in the results of their study where over 86% of users preferred their web interface to creating reminder lists, with only 6% using their cell phone interfaces. Another area

of differentiation is that in PlaceMail no information is extracted from the audio reminder, whereas in our system information extracted from intent lists serves the purpose of a “hint” about information contained in the audio. PlaceMail is also unable to support collaborative creation and access to lists, as their primary means of delivery is personal cell phones (a device that is typically not shared amongst users).

6 Conclusions

Shopping is a collaborative and social activity and tools that are successful in this domain need to extend beyond just capturing and sharing on the Web (like Google Notebook [3]). Since an impulse to buy can be created at any time and anywhere, and users may not have access to their PCs to capture this impulse, we believe that using multiple devices and integration of multiple modalities is needed to ease the process.

Our longitudinal study revealed that shopping is mostly goal-oriented. Users want to complete a certain goal (e.g. find an affordable birthday gift, buy food for potluck dinner), and they don’t necessarily have a particular structure of search in mind when they start. Thus the “highlight-and-click-to-add” feature of the web toolbar, and the availability of multiple modalities for adding items to the list provided a suitable degree of flexibility to the users and was extremely easy to learn.

Acknowledgments

We are grateful to all the participants of the user study. We would like to thank Kenton O’Hara for his insightful comments on our longitudinal study findings.

References

1. Froogle. Google product search, <http://www.google.com/products>
2. Gershman, A.V., McCarthy, J.F., Fano, A.E.: Situated Computing: Bridging the Gap between Intention and Action. In: Proc. ISWC 1999, pp. 3–9. IEEE Press, Los Alamitos (1999)
3. Google Notebook, <http://www.google.com/notebook>
4. Ludford, P., Frankowski, D., Reily, K., Wilms, K., Terveen, L.: Because I Carry My Cell Phone Anyway: Functional Location-Based Reminder Applications. In: Proc. CHI 2006, pp. 889–898. ACM Press, New York (2006)
5. Mankoff, J., Abowd, G.D.: Domisilica: Providing Ubiquitous Access to the Home. Technical Report GIT-GVU-97-17, GVU Center, Georgia Institute of Technology (1997)
6. Miller, D.: A Theory of Shopping. Cornell University Press, Ithaca (1998)
7. Sellen, A.J., Hyams, J., Eardley, R.: The everyday problems of working parents: implications for new technologies. Technical Report HPL-2004–37, Hewlett-Packard Labs (2004)
8. Thomas, A.: Grocery shopping: list and non-list usage. *Marketing Intelligence & Planning* 22(6), 623–635 (2004)
9. Underhill, P.: *Why We Buy: The Science Of Shopping*. Simon & Schuster, NY (2000)
10. Wolfenbarger, M., Gilly, M.: Shopping Online for Freedom, Control and Fun. *California Management Review* 43(2), 34–55 (2001)