

# From Opportunistic to Purposeful Information Discovery and Curation

Elena Voyloshnikova  
elenavoy@uvic.ca

Dr. Margaret-Anne Storey  
mstorey@uvic.ca

University of Victoria  
Victoria, BC, Canada

## Abstract

Everyday life revolves around the discovery and curation of digital information. People search the Web continuously, from quickly looking up the information needed to complete a task, to endlessly searching for inspiration and knowledge. A variety of studies have modelled information seeking strategies and characterized information seeking and curation activities on the Web. However, there is a lack of research on how existing Web applications support the discovery and management of information, especially concerning the motivations behind them and how different approaches can be compared.

In this paper, we present a study of information discovery tools and how they relate to the nature of information seeking. We propose a conceptual framework that deals with the opportunistic and purposeful aspects of how people discover and manage digital information. This framework can be used when designing new, evaluating, or updating existing Web applications to match expanded use case objectives.

## 1 Introduction

Today, people commonly use web technologies to support personal endeavors. These endeavors include a wide range of personal goals and activities, such as making decisions, planning projects, and

researching personal interests. For example, some people research their travel destinations using various online resources, and some shoppers look for product characteristics and conduct purchases online. In order to accommodate diverse and evolving user needs, web applications continuously introduce new features and services empowering goal-oriented activities.

Sometimes, web users aim to find particular pieces of information, such as showtimes and phone numbers, or perform particular actions, such as making purchases and banking transactions. Other times, users lack well-articulated objectives and information needs, which results in their actions, intentions, and information needs continuously changing and interweaving. Consequently, people often use web technologies to satisfy their personal interests adaptively or, in other words, opportunistically [ac].

Opportunistic use usually occurs when a person casually browses through sites in hopes to satisfy some mind curiosity, gain knowledge about a hobby-related topic, research long-term ambitions, or plan activities, events, and projects that she hopes to take on in the future [ac]. This mode of web use characterized by underdefined or absent information needs. For example, someone who is planning a hiking trip will try to research famous hiking trails in her geographic region, required physical fitness level, training strategies, hiking tips, etc. As noted above, the user is not looking for any specific information but rather trying to research general topics related to hiking.

Although opportunistic use can be related to

user's long-term goals, goal realization per se however, is not necessarily an outcome of this use. Often, users need to purposefully go online in order to acquire some information or perform an action that would help them achieve some outcome or complete some other activity. Therefore, purposeful use is characterized by a precise piece of information needed or an action required to complete some wider activity, such as making a decision or planning [ac]. For example, when a user plans to go to a concert, she might try to check if there is parking available by the concert venue. Here, parking availability is the missing piece of information that the user needs in order to make a transit-related decision.

It is important to note that purposeful use is not always associated with opportunistic. According to Lindley et al. [ac], this mode of web use occurs when users need to find some concrete information or perform an exact action, often in the context of some other wider activity. In this paper, we study purposeful web use only in the context of applications that support opportunistic doings. The fundamental assumption is that when users opportunistically search for information related to personal interests, they often save information for future exploitation and then use it for some concrete goal. For example, a person who browses the Internet to find a new book to read, might want to order that book online later on or look up its name so that she can put it on hold at the library.

To satisfy their information needs for diverse personal endeavors, people turn to various web technologies ranging from generic search tools to specialized web applications that provide relevant services. Several web usage studies have been directed at exploring high-level web tasks [ac], deriving models of information seeking behaviors [ac], and looking at methods of information curation [ac]. However, more research is necessary to determine how different tools and their features provide fundamental support for opportunistic use and further goal realization.

To enhance goal-related web-usage experiences and support users' high-level endeavors, we extend existing research by (1) deriving web application design elements for goal-oriented opportunistic web use, (2) validating those design elements by studying and describing currently existing web applications, and (3) providing guidelines to address elements of the framework when designing new ap-

plications. Specifically, our research goal is to gain understanding of how existing tools support goal-oriented opportunistic web use.

The remainder of this paper is organized as follows. Section 2 highlights some of the studies and technologies related to web usage. Section 3 outlines case study based methodology that was used for this study. In Section 4, we derive initial cognitive support elements for opportunistic mode of web use. Sections 5 and 6 deal with data collection and analysis phases of this research respectively. In Section 7, we describe the final framework, followed by limitations and the conclusion in sections 8 and 9.

## 2 Related Work

Several researchers have studied various characteristics of web usage behaviour. To gain understanding of how currently existing web tools support goal-oriented opportunistic web use, we study known mechanics of web usage, including high-level web tasks, information seeking behaviour, information curation, and modes of web use.

In [year] Kellar et al. [ac] separated web tasks into five categories: transactions, browsing, fact finding, information gathering, and other uncategorized tasks. In their later work, Kellar et al [ac] added communication and maintenance as additional web tasks. Similarly to Kellar, Sellen [ac] identified six web tasks that are commonly performed by web users: browsing, finding, house-keeping, information gathering, communicating, and transacting. Although Kellar et al. and Sellen make a clear distinction between different types of high-level tasks, it is evident that some information seeking tasks can overlap. For example, when gathering information about a new fitness studio, the user might be interested in its location. Location lookup is classified as fact finding; however, it contributes to the information gathering process and can be performed while the user is browsing.

At first, it seems convenient to think of user motivation behind various web tasks being conducting the task itself, such as communicating for the sake of communication. However, tasks such as browsing or fact finding can be motivated by higher goals, such as planning a weekend getaway or researching fitness programs and related topics. Therefore, when providing tool support for various tasks, it is useful to consider the motivation behind

these tasks, as it may differ. Morrison et al. [ca] makes a distinction between methods of web use and purposes. The authors derived a purpose taxonomy web use, including three purposes or motivations: finding information, comparing or choosing to make a decision, and using the web to find relevant information to gain understanding of some subject. Consequently, methods of finding information identified by Morrison et al. are collecting, finding, exploring, and monitoring. The differences between the two taxonomies suggest that different information seeking tasks may be performed to satisfy each of the information seeking purposes. Therefore, each purpose may require more than one task-supporting mechanisms.

A number of researchers have studied information seeking behaviour [ac]. Ellis et al [ac], proposed a model of information seeking characterized by six different patterns: starting, chaining, browsing, extracting, monitoring, and differentiating. Based on these patterns, Choo et al. [ac], derived corresponding anticipated web moves. According to authors [ac], when users identify sources of interest, they usually identify which web sites can point to that information of interest. Chaining occurs when users navigate through links on those initial pages. When people browse, they scan top-level pages, headings, lists, and site maps. Differentiating takes place when people bookmark, print, copy and paste information, or choose earlier selected site. Monitoring occurs when users revisit web pages or receive updates from some earlier visited sites. Finally, extracting can occur when the user systematically searches the site to extract information of interest.

In [insert year], Bates [ac] proposed a model of four information seeking modes that consists of being aware, monitoring, browsing, and searching. Bates differentiated the modes based on the levels of attention being active or passive, and information needs being directed or undirected. Thus, browsing can be characterized as undirected active information seeking because users do not know directly what information they are looking for, but they actively look for some information. Searching falls under active directed information seeking because the information need is clearly defined and the search is directed. Finally, monitoring and being aware are passive modes of information seeking although monitoring is directed, and being aware is undirected.

In [insert year], Bates [ac] extended her research with the notion of information farming. Information farming involves people collect and organize information for future use and revisitation. More commonly, it is referred to as digital curation, which is the notion of collecting and managing digital information for the purpose of adding value to the collection, and revisitation [ac]. Wiltaker [ac] believes that in terms of web use, there is happening a significant shift from information consumption to information curation, which means that people no longer just use the web to find and consume the information that they are interested in, including the information related to their personal endeavors, but they also try to save and manage that information so that later it can be reaccessed and exploited, at times for the purpose of goal realization.

Categorizing web usage into information seeking, digital curation, and other web tasks does not necessarily give full insight about how web usage related to personal goals is performed. Lindley et al. [ac] conducted a qualitative study involving 24 participants tracking their daily web usages in a form of a diary. As a result of this study, the researchers identified five distinct modes of web use: respite, orienting, opportunistic, purposeful, and lean-back [ac]. According to authors [ac], people web browse opportunistically when they look for information related to some personal interest, long-term goal, or future ambition. Purposeful use occurs when the user knows for a fact what information she needs to acquire or what action online she needs to perform in order to continue or finish some other activity. Respite mode usually occurs when users are in the process of waiting for something or taking a break, and it serves as means for people to temporarily occupy themselves when high engagement with the content is not a requirement. Orienting mode usually occurs when people want to be updated on what has been happening in their environment. Examples of this mode are checking email at work or looking at the news and updates on a social networking site. Finally, lean-back mode of web use can be thought of as listening to the radio or watching TV. It usually involves watching videos online or browsing through other types of entertainment content.

Lindley's et al. [ac] primary motivations behind looking at use modes that occur when people browse Internet was that traditional web use stud-

ies and web tasks discovered by other researchers could not reflect the depth of user's intentions online. It is evident that modes of web use make user's motivations clearer. In the context of goal-oriented activities, the mode that is best associated with user's long-term ambitions and personal hobbies and interests is opportunistic. Purposeful use often relates to immediate goals that feed into some wider context.

To further compare opportunistic and purposeful use modes, we summarized some essential differences between the two in Table 1. Opportunistic mode is usually performed because it relates to some user interest and the act of performing it is enjoyable by itself, whereas purposeful mode usually serves as a tool for completion or continuation of some other activity. An information need for opportunistic use can be blur and change continuously throughout web use session, whereas with purposeful use, people know exactly what information they are interested in and that information need usually does not change. Purposeful use has a well-defined point of completion, which occurs when objective is achieved or information is acquired, whereas with opportunistic use, people often cannot indicate completion of the web task and they finish whenever they have been browsing Internet for too long or they need to complete some other task of higher priority. Finally, opportunistic use is 'grasshopper-like': users jump from one resource to another. Purposeful use is characterized by direct actions or search. Motivation, information need, and duration, and usage patterns are the four characteristics that we will use in order to estimate whether a task is performed during the opportunistic mode of web use.

A number of researchers have studied how people use the web, including major tasks, information seeking behaviours, and information curation. However, there is lack of research on how currently existing technologies support people in performing these tasks in the context of goal-associated user intentions. With this overview of related work in mind, we proceed with building a framework of web application design elements for goal-oriented opportunistic web use.

### 3 Methodology

The study presented in this paper has two primary goals. The first goal is to build and validate a

framework of goal-oriented opportunistic web application design elements described in the following section. The second goal of the study is to gain perspectives on how to address different elements our framework when designing real-world applications.

In order to understand how applications support opportunistic web use, we first research how they support opportunistic information discovery (RQ1). Repeatedly mentioned by Lindley et al [ac], opportunistic use is often accompanied by users collecting and organizing discovered information, especially if the purpose of opportunistic use is to satisfy some long-term goal or ambition. Thus, in the second research question (RQ2), we ask about information curation support. Lastly, to learn about the goal-oriented aspect of existing opportunistic applications, we look at how opportunistic applications help with goal realization (RQ3). Therefore, the research questions are the following:

RQ1: How do existing web applications support opportunistic information discovery? RQ2: How do existing opportunistic web applications support information curation? RQ3: How do existing opportunistic web applications support immediate goal realization?

Our methodology for studying existing web applications is based on Yin's guidelines for designing a case study [ac]. The motivation behind choosing a case study over other methods of qualitative research was based on our choice of research questions, which have explanatory nature, lack of control over existing applications and their development, and having to focus on contemporary use of real-life web applications. According to Yin [ac], case study would be an optimal research strategy given above characteristics of the subject matter.

To answer each of the research questions, we conducted three independent case studies, where we considered web applications as our units of analysis, or 'cases'. For each case, we examined overall purpose of the application, its description as defined by the site itself, and available literature and documentation related to the case against the features present at a site. For example, if an application has bookmarking capabilities as well as means for information categorization (features), we would check if it is indeed used for information curation purposes (site description, literature, and documentation).

To increase external validity of our study, we

chose the cases based on replication logic [ac]. Using replication logic in a case study design means carefully selecting each case so that it either predicts analogous results or predicts contrasting results but for anticipated reasons [ac]. Choosing cases to predict analogous results is called literal replication, whereas choosing cases to predict contrasting results is called theoretical replication. Carefully following this logic allows for analytic generalization when it comes to generalizing results of the study [ac].

Before conducting the full study, we derived a set of initial theoretical propositions that would let us predict the results of analyzing each case. Therefore, each case study had a form of an iterative process of selecting cases, analyzing each case, and determining whether or not each case meets the theoretical propositions. If it did not support the theory, then we modified the propositions and selected a new set of cases until the results of analyzing the case gave the anticipated results for all cases. We then transformed the final set of propositions into a set of design requirements and constructed a goal-supporting opportunistic framework (see Fig. 1, with design elements E1-E12) by grouping and connecting the design elements. The following subsections provide more details about each chosen case and logic behind choosing each case for each study. Limitations of our study are presented in Section 6.

### 3.1 Case Study 1: Opportunistic Information Discovery

The first case study we conducted aimed at answering our first research question: RQ1: How do existing web applications support opportunistic information discovery?

Based on preliminary research and continuous iterations of this case study, we formulate the first proposition: a web application strongly supports opportunistic information discovery if it provides information access by bringing insightful information together in a uniform way and integrates with the global information space; supports browsing activities by allowing the user to directly navigate to resources related to the topic of interest and allows for undirected navigation among those resources; and supports visual information representation by providing visual preview of the resource and arranging resources in a spatially meaningful

way.

We used nine final cases (C1.1-C1.9, see Table 2) for this study. Cases C1.1-C1.3 were used for literal replication, meaning that the cases supported opportunistic information discovery as described in our proposition. Cases C1.4-C1.9 were used for theoretical replication, with C1.4 and C1.5 lacking information access capabilities and C1.6 and C1.7 lacking topic-oriented navigation features and having lack of insightful information with the content being primarily selected for entertainment purposes. Finally, C1.8 and C1.9 have poor visual representation, and therefore, lack some support for opportunistic information discovery.

### 3.2 Case Study 2: Information Curation Within Opportunistic Applications

The second case study was designed to answer RQ2: RQ2: How do existing opportunistic web applications support information curation?

The final theoretical proposition can be stated as follows: an opportunistic web application strongly supports information curation if it provides bookmarking capabilities, means for custom information categorization and social information curation. We used five final cases (C2.1-C2.5, see Table 3) for this study, all of which were opportunistic web applications. Cases C2.1-C2.2 were used for literal replication, meaning that the cases supported information curation by providing means for bookmarking and custom information categorization, and allowing for public information sharing. Cases C2.3-C2.5 were used for theoretical replication, meaning that the cases did not support information curation because they lack design elements mentioned above.

### 3.3 4.2. Case Study 3: Goal Realization Within Opportunistic Applications

The third case study was designed to answer RQ3: RQ3: How do existing opportunistic web applications support immediate goal realization?

The final proposition can be stated as follows: an opportunistic web application strongly aids goal realization if it supports revisitation mechanisms, and provides specific information and means for associated actions and transactions We used six fi-

nal cases (C3.1-C3.6, see Table 4) for this study, all of which were opportunistic web applications. Cases C3.1-C3.3 were used for literal replication, meaning that the cases supported goal realization by empowering purposeful use using with design elements from proposition above. Cases C3.4-C3.5 were used for theoretical replication, meaning that the cases do not aid goal realization effectively because they did not allow for purposeful use, such as looking up specific information or performing specific actions. C3.6 was chosen because, as many other blogs, it lacks revisitation mechanisms.

## 4 Conceptual Framework

Although information discovery and curation tasks elicit predominant portion of interactions within many Web applications today, there is lack of systematic guidelines for designing and evaluating applications that could afford these tasks. We hope to reduce this gap by developing a framework of design factors that enable digital information discovery and curation.

The framework (see Table 1) consists of two main categories of factors, discovery and curation, that are consequently broken down into subcategories. Each subcategory contains factors that determine use case enablers and corresponding questions to ask when designing or evaluating an application. This section outlines main components of the framework.

### 4.1 Information Discovery

A number of researchers attempted classification of information seeking tasks and methods. In our framework, we built on existing research to derive corresponding design factors. Thus, the discovery category consists of serendipitous discovery, fact discovery, rediscovery, and channel-based discovery.

#### 4.1.1 Serendipitous discovery

Serendipitous discovery refers to information discovery resulting from serendipitous browsing. Such discovery is characterized by underdefined, absent, or hidden information need, and it is usually involves browsing through diverse resources with varying content types. The following is the key criteria that influence serendipitous information discovery:

- **Arbitrary navigation.** In order to browse through
- **Search-based navigation.**
- **Navigation with a dictionary.**
- **Integration.**
- **Visual preview.**
- **Spatial arrangement.**

#### 4.1.2 Fact discovery

- **Arbitrary navigation.**
- **Search-based navigation.**
- **Navigation with a dictionary.**

#### 4.1.3 Rediscovery

- **History-based rediscovery.**
- **Bookmark-based rediscovery.**
- **Search-based rediscovery.**

#### 4.1.4 Subscription-based discovery

A number of researchers have talked about subscription-based information distribution []. However, Although blog subscriptions have been around before microblogging gained popularity, channel based discovery significantly

- **Site subscription.**
- **User subscription.**
- **Notifications.**
- **News stream.**

### 4.2 Information Curation

#### 4.2.1 Management

- **Categorization.**

#### 4.2.2 Information Preservation

- **Internal Preservation.**
- **External Preservation.**

#### 4.2.3 Augmentation

- **Evaluation.**
- **Annotation.**

Table 1: Conceptual Framework.

Factors	Questions
<b>Discovery</b>  <i>Serendipitous discovery</i> Arbitrary navigation  Search-based navigation  Navigation with a dictionary  Integration  Visual preview Spatial arrangement  <i>Fact discovery</i> Search-based navigation Navigation with a dictionary Uniform representation  <i>Rediscovery</i> History-based rediscovery Bookmark-based rediscovery Search-based rediscovery  <i>Subscription-based discovery</i> Site subscription User subscription Notifications News stream	Does the application provide means for arbitrary navigation among resources? Does the search feature help discover arbitrary resources related to the topic of interest? Does the dictionary suggest and help navigating to resources related to the topic of interest? Does the application support navigation to information sources beyond the application? Do resources have visual previews? Are resources presented in a spatially meaningful way?  Does the search feature help discover the specific resource of interest? Does the dictionary help narrow results to specific types of resources? Are resources presented in a uniform way?  Does the application save and provide access to browsing history? Does the application support bookmark-based resource revisitation? Is the search within application a reliable method for resource revisitation?  Does the application allow subscriptions to news and updates? Does the application allow subscriptions to other users' activities? Does the application have notification mechanism(s)? Can subscription updates be visible within the application?
<b>Curation</b>  <i>Management</i> Categorization  <i>Preservation</i> Internal Preservation  External Preservation  <i>Augmentation</i> Evaluation Annotation  <i>Social curation</i> Adding resources  Resharing resources	Does the application support information categorization?  Does the application have bookmarking mechanism for preserving information within the application? Does the application have bookmarking mechanism(s) for preserving information outside of the application?  Can resources be evaluated? Can resources be annotated?  Can resources be added to the publicly available pool of information from outside of the application? Can resources be publicly reshared within the application?

#### 4.2.4 Social Curation

Recently, a new trend in bookmarking, social bookmarking, started to gain popularity. Social bookmarking refers to the notion of bookmarking and organizing pages and other resources on the web and sharing them with other people by means of social bookmarking systems [ac]. In other words, social bookmarking is a commonly accepted synonym of social information curation. One of the first visions of social bookmarking was associated with web blogging. Oravec [ac] believed that web blogs help “annotate” or “bookmark” important information and build a “map” of the Internet. However, blogs usually have a limited number of moderators that curate information within the blog.

- **Adding resources.**
- **Resharing resources.**

## 5 Discussion

Through analysis of twenty different applications in total, we were able to derive twelve key design elements that can help support goal-oriented opportunistic use. However, depending on the nature of the information need and the possible goals behind researching a topic, the amount of support needed can vary. Thus, there exists a large spectrum of goal-supporting opportunistic applications. Some have no support for information curation, others have no means for goal realization or visual information discovery, etc.

We do know for a fact that facilitating information curation is not the determinant of an application being goal-oriented or opportunistic. Blogs, for instance, provide no formal means for information preservation and management. However, they can surely be browsed opportunistically [ac], and often, people can accomplish their goals depending on their nature of what they would like to learn about. Blogs also do not provide the same social curation capabilities as truly social-networking sites.

Analogously, the amount of support needed for accomplishing a goal can vary. On some occasions, the information need is limited to visual information, then no other specific information is required. Sometimes, there are no associated actions of transactions that can be performed. If the purpose of opportunistic information discovery can be satisfied

within a single browsing session, then no revisitation mechanisms are needed. Similarly to information curation, goal realization design elements cannot be considered as essential requirements of an opportunistic web application. Although they help goal-realization, goal realization itself is not the final target of opportunistic use.

It is challenging to say for a fact what are the bounds of the framework’s design elements’ implementations. It is more probable that the bounds do not even exist. All of the methods of implementing each of the design elements are simply examples of how real-world applications use them. Spatial information representation, for example, varies even within some applications (Pinterest, Youtube). Visual preview is often limited because of the resources that can be previewed.

Applying the framework can extend existing tools’ current uses. For example, Google Maps effectively supports purposeful use by letting users search for places and their addresses. Making information more visual and rearranging information displayed, as well as suggesting categories to facilitate opportunistic information discovery can expand the tools use from just purposeful to opportunistic. Letting the users extensively curate information that they might be interested in would not only add value to the information that is already available through the Google Maps, but would also facilitate engagement.

As another example, Facebook is known to be used for socializing or networking purposes. Applications such as Pinvolve, extend its use and make it possible to discover information in a visual format as well bookmark it and categorize. Thus, the purpose of facebook is extended from being used purposefully, in respite, lean-back, and oriented modes, to being used in opportunistic goal-oriented mode.

Currently existing ecommerce environments often have mechanisms similar to what one can find in a goal-oriented opportunistic web application. However, they often lack information curation (especially social information) curation mechanisms. A lot of online stores, however, utilize the social and curation capabilities of other applications, such as Pinterest and PolyVore to sell their items. When people republish different items, they often categorize them, add comments, description, or just like them. All of those aspects of information curation add value to the item, so it can become more ap-



peeling for buyers to purchase.

Another application of our framework is to describe existing tools in the context of goal-oriented opportunistic doings.

## **6 Limitations and Threats to Validity**

## **7 Future Work**

One of the possible future research objectives would be to test the framework on a real-world application, and to either enhance its use as an opportunistic goal-oriented application, or to extend its use to support opportunistic information discovery, information curation, or goal realization.

## **8 Conclusion**

## **References**

- [1] Beagrie, Neil. "Digital curation for science, digital libraries, and individuals." *International Journal of Digital Curation* 1.1 (2008): 3-16.
- [2] Java, Akshay, et al. "Feeds That Matter: A Study of Bloglines Subscriptions." *ICWSM*. 2007.