

Forward, Back and Home Again

**Analyzing User Behavior
on the Web**

Eelco Herder

Abstract

In a period of less than two decades, the World Wide Web has evolved into one of the most important sources of information and services. Due to the infancy of the Web and its rapid growth, our knowledge on how users interact with the Web is limited - knowledge which is likely to provide pointers for improvements in the design of Web sites and Web browsers. In this thesis, we aim to provide an integrative overview of theoretical insights and empirical findings, and to extend this body of knowledge with results from a number of user studies.

Research in the fields of user modeling and adaptive hypermedia, and the sub domain of Web personalization, is specifically aimed at understanding and supporting individual users' interests, needs, and interaction patterns. From the literature overview it becomes apparent that these techniques are useful in limited contexts, yet it is not evident whether general Web interaction carries sufficient evidence in order for these techniques to be useful. A large amount of Web activities has been shown to be associated with *recurrent tasks*; unfortunately, most data is limited to statistics that do not explain why and how users revisit pages. A further significant Web activity is *information gathering*. Theoretical models characterize this process of users following their *information scent*; empirical studies indicate that information gathering often involves *orienteering behavior* - an interleaving process of issuing queries and gathering context by navigating. A common usability problem in both activities is termed *disorientation*, which refers to situations in which users fail to understand their current position in a Web site, the way that led to this position, and where they can go to. A number of studies were carried out to shed a new light on the above matters. An analysis framework was developed for this purpose, based on first-order *navigation measures* and second-order *navigation styles*, *visualization* of user navigation paths, and the *cleaning* of client-side usage logs.

In our laboratory studies we targeted the issue of disorientation. Conflicting results exist on what interactions patterns indicate that users may experience such usability issues. In the first study we identified two aggregated measures

- navigation styles - that were closely associated with perceived disorientation. *Flimsy navigation* is a weak navigation style, mainly exhibited by inexperienced users who appear not to be able to reconstruct their navigation paths, and therefore are prone to get stuck. Task-related structural navigation support for both forward navigation and backtracking is expected to help users who navigate in a flimsy manner. By contrast, *laborious navigation* is an exploratory navigation style that emerges at the beginning of a task and serve to quickly build a mental model of the site structure. We believe that theoretical models of Web navigation should consider laborious *orienteering* as enrichment activities rather than failure, as it may result in increased performance at a later stage. The results from the first study were confirmed by the second laboratory study, in which we investigated the impact of task-based *link suggestions* on user navigation styles. From the results it became clear that users who were provided with this type of adaptive navigation support, navigated in a more structured manner, with more targeted backtracking and less returns to the sites' home pages. The link suggestions were positively received, improved user perceptions, and decreased task execution time.

From earlier user studies it is known that users often return to Web pages that they have visited before. In a long-term study, we aimed to distinguish the various kinds of revisits, and to find out to what extent browser support for these revisits may be improved. For this purpose, we tracked and analyzed the Web activities of 25 participants for an average period of three months. We distinguished cross-session revisits (*recurrent behavior*) from within-session revisits (*backtracking*) - the latter activity covered 74% of all revisits. Whereas backtracking is surprisingly well supported by a list of the 15 most recent pages, a list to cover a reasonable percentage of recurrent activities would need to be overly long. In particular infrequent and irregular page revisits are problematic, as the user is less likely to remember the exact addresses; ironically, it is also less likely that the address still resides in the browser history.

In contrast to the earlier studies, our participants often made use of multiple browser windows or tabs. This interaction style reduces the need for backtracking, and allows for side-by-side comparison of results. A disturbing consequence is that the concept of the back button is disrupted; as the history is split over several stacks, users need to keep track of what activities were carried out in which window; this places a significant cognitive load on the user. A further change is the highly increased percentage of form submissions, which indicates that the Web has become a hybrid between 'traditional' hypermedia and a dynamic, interactive system. This requires a radical rethinking of the browser interface.

From the results several directions for navigation support can be extracted. Based on the orienteering behavior that we observed in our studies, it is likely that navigation will remain an important paradigm in Web-based interaction; the combined process of search and exploration of information patches allows users to find what they need rather than a best match to some query. Currently, search

and navigation are largely separated activities, and users need to actively keep track of their query refinement process. A visual query modification history, link relevancy indicators, and search result ranking that takes surrounding pages into account, are expected to be simple yet effective means to support this process.

As users often make use of multiple windows, users need to keep track of what activities were carried out in which window; branched history mechanisms that takes both the temporal order of user navigation, and the parallel character of multiple windows into account, are mechanisms that definitely need to be developed. In addition, results from our laboratory studies indicate that annotation of recently followed paths and links that lead to recently visited pages may greatly reduce user disorientation.

Users often carry out recurrent activities on the Web. Whereas the larger part of these activities involve a small number of very popular pages, the greatest need of history support for these activities involves a large number of infrequently visited pages; the number is too large to be covered by a reasonably small history list or bookmark file. As an alternative, we propose explicit history search mechanisms, combined with visually annotated trails from earlier search results - which are often waypoints rather than the desired item to be revisited.

Perhaps the most important challenge for current Web browsers is the shift of the Web paradigm from a hypermedia system to a hybrid of hypermedia and interactive applications. Document management options are dearly needed to deal with volatile yet relevant dynamically generated pages, such as travel plans and order confirmations. Ideally, users should only need to access one point both for revisiting earlier travel plans, and for creating new ones. This will require a major rethinking of the metaphor of Web navigation into an interleaving process of navigation and interaction.

Several theoretical models and empirical studies have provided us with various perspectives on how users interact with the Web. It has become clear that there is no such thing as the 'typical' Web user or a 'typical' Web navigation session. It is most unlikely that the various aspects can be covered by one integrated model. Therefore, we think it is more desirable to strive toward a comprehensive body of empirical data and theoretical insights, with explicit indications in which situations they are applicable.

Personalization of the Web interface to the individual user has proven to be effective in limited environments. Given the sheer variety of tasks for which the Web is used, sophisticated user-adaptive mechanisms for general Web navigation are likely to be infeasible. Instead, we think a current challenge for the adaptive hypermedia community is to develop and evaluate straightforward support mechanisms, which may well turn out to be one-size-fits-all designs for diversity.