Catching the User - User Context through Live Logging in DAFFODIL

Claus-Peter Klas
FernUniversität in Hagen
claus-peter.klas@fernuni-hagen.de

Matthias Hemmje
FernUniversität in Hagen
matthias.hemmje@fernuni-hagen.de

ABSTRACT

This demo will present the logging facilities to capture the user context within the Daffodil framework during a live search in computer science data sources. We propose to use the Daffodil system as an experimental framework for the evaluation and research of interactive IR. The system already provides a rich set of working services and available information sources. These services and sources can be used as a foundation for further research going beyond basic functionalities. In addition, the system can easily be extended regarding both services and sources. Daffodil's highly flexible and extensible agent-based architecture allows for easy integration of additional components and access to all existing services. Finally, the system provides a user-friendly graphical interface and facilitating services for log generation and analysis. The experimental framework can serve as a joint theoretical and practical platform for the evaluation of DLs, with the long-term goal of creating a community centered on interactive IR and DL evaluation.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**]: User Interface, Logging; H.3.7 [**Digital Libraries**]: User Issues

General Terms

information retrieval, visualization, interactive systems

1. INTRODUCTION

It is our intention to demonstrate the Daffodil system with regard to capturing the user context by logging all user initiated actions. Due to the rich functionality of the DAFFODIL system the user is able to explore the search domain in a comprehensive and sustainable way. During the interaction of the user with the system we capture the behavior through logging the actions. We categorized these actions into ten different conceptual events described in [1].

1.1 Daffodil-Framework

DAFFODIL consists of several tools to support the users search tasks. We would like to demonstrate the DAFFODIL [1] framework as an experimental system for the evaluation of the interaction in information search and retrieval. DAFFODIL is a virtual digital library system providing access to many sources from the domain of computer science, and targeted at strategic support of users during the information seeking and retrieval process. It provides basic and high-level search functions for exploring and managing information objects including annotations over a federation of heterogeneous digital libraries (DLs) based on a service-oriented architecture. For structuring the functionality, we employ the concept of high-level search activities for strategic support. A comprehensive evaluation in [2] showed that the system supported most of the information seeking and retrieval aspects needed for a computer scientist's daily routine.

Additionally DAFFODIL incorporates the concepts of adaptivity [3], collaboration, recommendation and awareness. In order to enable adaptivity and recommendations DAFFODIL collects implicit and explicit user interactions and system actions as described in previous publications [1]. This interaction can be examined and captured at various levels of abstraction, starting at the system/hardware level and covering the complete spectrum of user-system interaction.

2. LEVELS OF LOGGING

When using transaction logs for evaluation, the primary levels surveyed are the **user**, the **system**, and the **content** that is being searched, read, manipulated, or created. Because interaction between the system and the user can be captured at various levels of abstraction we focus on three levels of evaluation:

User behavior level: Data about users and their behavior are located at this level. Each user has a task to accomplish, within a certain social environment, and brings her individual knowledge to that task.
Concept level for comparative evaluations: The concept level captures data about generalized events generated by the DL user. By logging these events, user evaluation can be backed up with statistical data and a comparative evaluation of different users, systems and system content can be undertaken.
System level: System events happen on the computer or in the computer network where DL services are executed. This

[1] `http://www.daffodil.de`

level aggregates specific information concerning the state of the DL (e.g., database conditions, server load, or amount of network traffic) and its response (e.g., response time).

Through logging events by level, we have a horizontal view that tracks a sequence of events dealing with a single aspect of the DL. For example, by focusing on events that occur on the concept level, we can identify the user's moves and tactics as she works her way through the document space. In contrast, a vertical view across levels gives us information about the impact of a specific event across the DL system, from information about user behavior on the highest level to system specific data on the lowest level.

3. EVENTS ON THE CONCEPT LEVEL

At the concept level, we have identified several general event types that support comparative evaluation across DLs. Our focus on the concept level represents the centrality of these events for log analysis and interpretation: events that occur on the concept level indicate critical aspects of the user's interaction with the DL system and supply valuable data for rich interpretation of user behavior. As is highlighted in other DL logging studies [4], current approaches are often inadequate for capturing complex or abstract actions by the user and are therefore unable to elicit meaningful conclusions. By logging data about general event types at the concept level, we provide a basis for comparative evaluation across DLs and still gain insides into the users behavior.

The event types and event properties that we have identified are neither fixed nor complete and should be viewed as recommendations. that can also serve as discussion points in the community.

We have identified the following events on the concept level:
Search: The user formulates a *query* or *filter condition* that is to be processed by a given DL service against a *collection*.
Navigate: The user initiates an event by selecting from a set of possible moves from one point to another.
Inspect: The user accesses the details of a single object.
Display: The display event describes a specific *visualization* of the information presented to the user.
Browse: The user selects an event for viewing a set of DL objects (e.g., viewing a result list following a search).
Store: The user files an object for later reuse, either at a generic location (e.g., on a clipboard) or at a specific location (e.g., in a specific folder).
Annotate: The user adds information to an existing DL object, either by marking specific parts of it, by linking it to other digital library objects, or by adding inline or external comments.
Author: The user creates a new DL object or edits an existing object such as a document or annotation.
Help: The user requests help or information. The help event may be general or context-specific and can include introductory overviews or tutorials about the DL system.
Communicate: Users collaborate through communication, either by posing a simple question or through the use of specific tools or services.

4. DEMONSTRATION

In the demonstration we will present the DAFFODIL system with a specific focus on the live logging of user events.

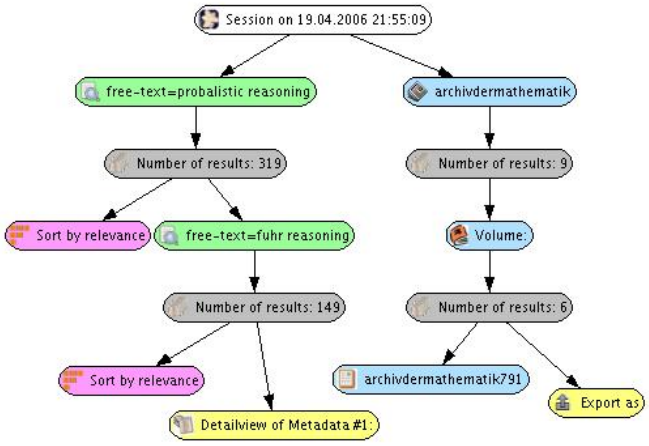

Figure 1: top-down tree visualization

Based on a given task several search and browse tools will be presented and a analysis, e.g. for relevance feedback will be given. In figure 1 a small search session if graphically presented as tree visualization. The colors of the nodes in the figure correspond to the concept level.events for easier recognition.

This captured information represents the basis to further understand and support the user. It of course does not excuse from running a real user evaluation. Such support could be done through recommendation via implicit relevance feedback as well as collaborative recommendations through other users in a similar situation. We think, that given the context model within the Daffodil-Framework, we are able to understand and categories user behavior and provide solid data to support system oriented IR evaluation, e.g. based on user simulation.

5. ACKNOWLEDGMENTS

This work is funded via the German Science Foundation (DFG) in the project LACOSTIR.

6. REFERENCES

[1] C.-P. Klas, H. Albrechtsen, N. Fuhr, P. Hansen, S. Kapidakis, L. Kovács, S. Kriewel, A. Micsik, C. Papatheodorou, G. Tsakonas, and E. Jacob. A logging scheme for comparative digital library evaluation. In *ECDL 2006*, Lecture Notes in Computer Science, pages 267–278, Heidelberg et al., September 2006. Springer.

[2] C.-P. Klas, N. Fuhr, and A. Schaefer. Evaluating strategic support for information access in the DAFFODIL system. In *ECDL 2004*, Lecture Notes in Computer Science, Heidelberg et al., 2004. Springer.

[3] C.-P. Klas, S. Kriewel, and M. Hemmje. An experimental system for adaptive services in information retrieval. In *Proceedings of the 2nd International Workshop on Adaptive Information Retrieval (AIR 2008)*, October 2008.

[4] B. Pan. Capturing users behavior in the national science digital library (nsdl). Technical report, NSDL, 2003. `http://dlist.sir.arizona.edu/848/`.