

DAFFODIL – Nutzerorientiertes Zugangssystem für heterogene Digitale Bibliotheken¹

Claus-Peter Klas und Sascha Kriewel und Norbert Fuhr und Andre Schaefer

Zusammenfassung

DAFFODIL ist ein System zur integrierten Suche in verteilten, heterogenen Digitalen Bibliotheken mit der Zielsetzung der strategischen Unterstützung des Benutzers. In diesem Artikel wird ein Überblick bezüglich des aktuellen Standes der Funktionalität gegeben.

Einleitung

In zahlreichen Fachgebieten wird heutzutage eine große Anzahl Digitaler Bibliotheken zur Literaturrecherche genutzt. Die föderierte und in Funktionalität, Anfragesprache und Benutzeroberfläche heterogene Natur der Angebote macht Systeme notwendig, die einen einheitlichen Zugriff bieten. Ein solcher Zugriff auf föderierte Bibliotheken ist jedoch für viele Anwendungen nicht ausreichend, sondern es sind reichhaltige und höhere Suchfunktionalitäten notwendig, um Recherche auf effektive Art zu unterstützen (siehe Fuhr et al. (2000)).

Hier bietet das System DAFFODIL eine integrierte Suche in heterogenen Digitalen Bibliotheken eines Fachgebiets unter Zusammenführung der Ergebnisse. Zur Zeit existiert eine prototypische Realisierung für den Bereich der Informatik, welche die Suche in mehr als zehn verschiedenen Digitalen Bibliotheken und anderen Informationsquellen ermöglicht. Darüber hinaus erfahren die Benutzer bei ihrer Suche strategische Unterstützung durch die angebotenen höheren Suchfunktionen, umgesetzt in einer Reihe integrierter Werkzeuge (siehe Fuhr et al (2002)).

DAFFODIL verbindet in seiner Oberfläche in natürlicher Weise Browsing- und Suchstrategien. Dabei unterstützt das System die kollaborative Recherche und bietet *Awareness*, d.h. es wird Information über neue oder geänderte Objekte bezüglich früherer Recherchen geliefert. Die Benutzer sind dabei frei in der Wahl einer Strategie für ihre Suche.

Zudem erfährt der Benutzer Unterstützung bei der Durchführung von beliebigen Suchtaktiken oder -strategemen und bei der Kombinationen solcher, die u. a. zu umfangreicheren Suchplänen führen können. Bates (siehe Bates (1979), Bates (1990)) hat, gestützt auf empirische Beobachtung, eine Reihe erfolgreicher Taktiken für die Informationssuche identifiziert. Im Folgenden wird auf die aktuelle grafische Oberfläche einschließlich der wesentlichen Werkzeuge eingegangen. Eine kurze Darstellung der Ergebnisse einer Evaluation der strategischen Unterstützung der Benutzer und der angebotenen pro-aktiven Funktionen folgt im Anschluss daran.

Der Beitrag schließt mit einem kurzen Ausblick ab.

¹ Das Projekt DAFFODIL wurde gefördert durch die (DFG) innerhalb des Schwerpunktprogramms „Verteilte Verarbeitung und Vermittlung Digitaler Dokumente“ V3D2

Graphische Oberfläche und ihre Werkzeuge in DAFFODIL

Die Auswahl der im Folgenden vorgestellten Werkzeuge konzentriert sich auf die wesentlichsten Dienste, die den Benutzer dieses Systems mit höheren Suchfunktionalitäten unterstützen. Der DAFFODIL-Desktop (siehe Abbildung 1: Desktop: Suchwerkzeug, Persönliche Handbibliothek, Verwandte Begriffe, Autorennamen-Vervollständigung) ist auf einfache Erweiterbarkeit ausgelegt, die Integration von weiteren Werkzeugen ist daher leicht durchzuführen.

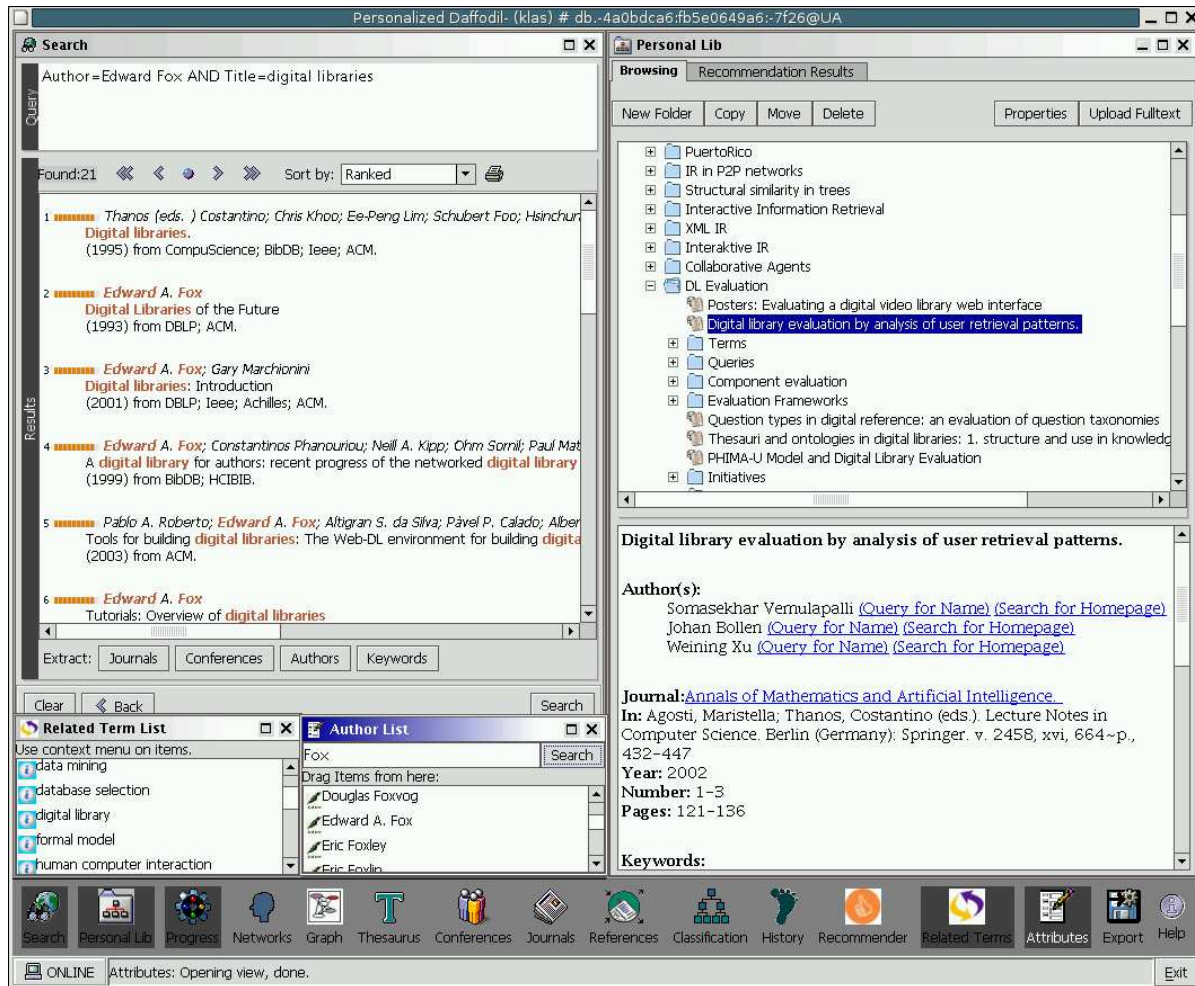


Abbildung 1: Desktop: Suchwerkzeug, Persönliche Handbibliothek, Verwandte Begriffe, Autorennamen-Vervollständigung

Das Suchwerkzeug

Das Suchwerkzeug, dargestellt in Abbildung 2: Suchwerkzeug: Eingabemaske zur Suche in Digitalen Bibliotheken mit Auswahl der Datenquellen, stellt einen der möglichen Ausgangspunkte einer Literaturrecherche in DAFFODIL dar. Es bietet eine für den Benutzer einfach zu handhabende Eingabemaske, die es erlaubt, einheitliche Anfragen an die verteilten Bibliotheken zu formulieren und dadurch die Suchdomäne durch Auswahl einiger oder aller digitalen Bibliotheken zu spezifizieren.

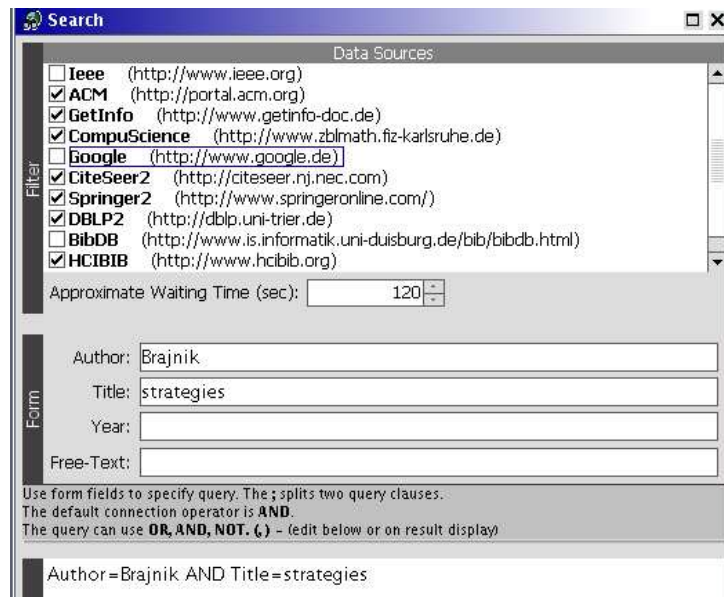


Abbildung 2: Suchwerkzeug: Eingabemaske zur Suche in Digitalen Bibliotheken mit Auswahl der Datenquellen

Die Anfragen des Suchwerkzeuges werden über Agenten und Wrapper an die Informationsanbieter weitergeleitet. Diese Wrapper nutzen hierzu in der Regel das WWW-Suchformular der entsprechenden Digitalen Bibliothek, die durch das eigens für DAFFODIL entwickelte Wrappertoolkit, gestützt auf XML, XPath und regulären Ausdrücken, mittels Konfigurationsdateien die Anfrage- und Resultattransformation durchführen. Die Anfragen werden parallel bearbeitet, die Ergebnisse werden anschließend zusammengeführt und dem Benutzer im derzeitigen Systemstatus auf homogene Weise zur Betrachtung und Navigation in Form einer gewichteten Resultatliste präsentiert. Dabei werden über unaufdringliche Ikonen (siehe Abbildung 3: Gewichtete Resultatliste mit Ikonen) bereits betrachtete, abgespeicherte oder anderweitig bearbeitete Dokumente kenntlich gemacht.

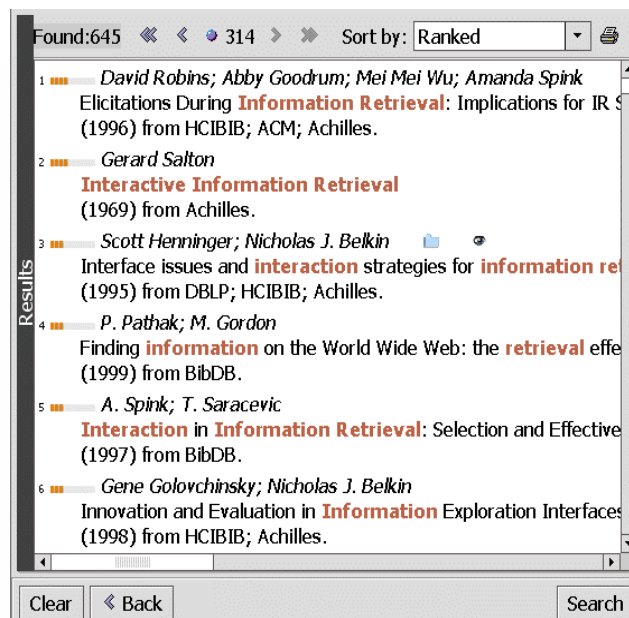


Abbildung 3: Gewichtete Resultatliste mit Ikonen

Mit Hilfe einer Extraktionsfunktion kann dem Benutzer ein aus der Resultatliste gewonnener Überblick über die zentralen Autoren, Konferenzen, Journale oder Begrifflichkeiten einer Suchergebnisse geboten werden. Dadurch wird z. B. das Auffinden von Hauptakteuren eines Themengebietes erleichtert und zudem stehen dem Benutzer neue Ansatzpunkte für die weitere Suche zur Verfügung.

Persönliche Handbibliothek

In der persönlichen Handbibliothek (siehe Abbildung 4) ist das Speichern von Ergebnissen, Volltexten, Stichworten oder Suchtermen, Autoren, Konferenzen oder Journalen über den Kontext einer Suche hinaus möglich. In persönlichen Ordnern können Benutzer Ergebnisse strukturiert ablegen und so über mehrere Suchsitzungen hinweg ein Archiv ihrer Literaturrecherchen aufbauen. Für die Benutzung der Resultate außerhalb des Systems stehen Exportfunktionalitäten zur Verfügung. In so genannten Gruppenordnern besteht ebenfalls die Möglichkeit zur Annotation an Objekten, einschließlich bestehender Annotationen, wodurch das Konzept kollaborativer Recherche unterstützt wird. Auf Grundlage der persönlichen Ordner bietet das System *Awareness* und Empfehlungen.

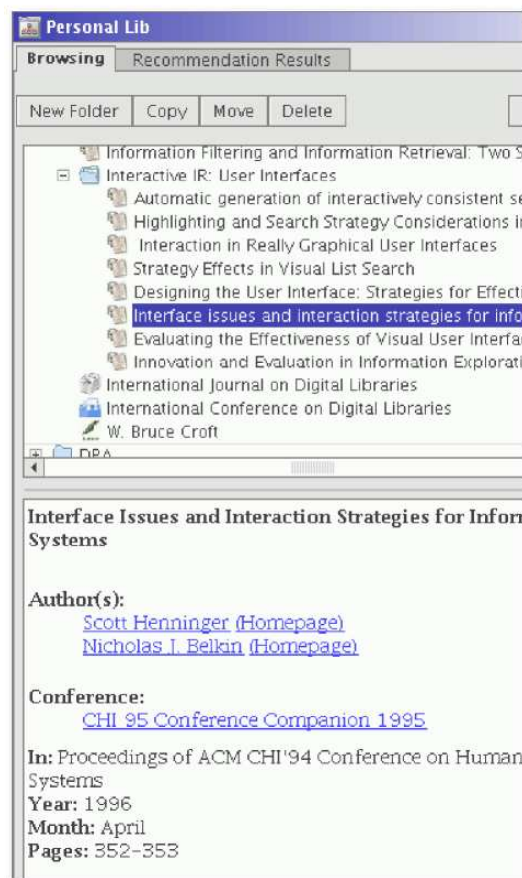


Abbildung 4: Persönliche Handbibliothek: Strukturierte Ablage aller digitalen Bibliotheks Objekte

Um ein Thema über einen längeren Zeitraum bearbeiten und verfolgen zu können, kann man derzeit die Objekte Autor, Suchanfrage, Journal, Konferenz und Webseite mit dem Konzept beobachten und sich durch einen *Notification Service* über neue Resultate informieren lassen. Dies beinhaltet eine automatische, systemseitige Suche nach neuer Information in festgelegten

Intervallen. Die so neu gefundenen Ergebnisse werden in die persönliche Handbibliothek eingestellt und kenntlich gemacht.

Strategische Unterstützung

Während herkömmliche Suchsysteme oft nur einfache, grundlegende Suchoperationen unterstützen, wird von DAFFODIL das Ziel verfolgt, durch die enge Integration von höheren Suchfunktionen gerade Nichtexperten als Benutzern eine effektivere Suche zu ermöglichen.

Bates (1990) hat in diesem Zusammenhang drei höhere Abstraktionsebenen für die Kategorisierung von Suchfunktionalitäten angeführt:

- *Taktiken* sind als eine oder mehrere kombiniert eingesetzte, grundlegende Suchoperationen zu verstehen. Damit sind beispielsweise die Schritte gemeint, die notwendig sind, um eine Suchanfrage zu verfeinern oder einen Suchbegriffs zu generalisieren.
- *Strategeme* stellen eine komplexe Abfolge von Aktionen aus Grundoperationen und Taktiken dar, welche die Strukturen einer Domäne zur Suche ausnutzen.
- Unter einer *Strategie* soll dann ein vollständiger Plan zur Ausführung einer Recherche verstanden werden. Strategien setzen sich im Allgemeinen aus einer Reihe von Grundoperationen, Taktiken und Strategemen zusammen.

DAFFODIL beginnt auf der Ebene der Strategeme und bietet durch strategische Unterstützung Benutzern die Möglichkeit, komplexe Suchstrategien zu verfolgen. Zusätzlich werden die Strategeme innerhalb von DAFFODIL in fünf verschiedene Phasen nach Weibel und Miller (1997) und Paepcke (1996) – *Discover, Retrieve, Collate, Interpret* und *Re-Present* – differenziert, um einerseits die Komplexität des Systems zu verringern und andererseits den Benutzer in jeder dieser Phase gezielt zu unterstützen.

- **Discover:** In dieser Phase steht die Suche und Auswahl der Datenquellen durch den Benutzer im Vordergrund.
- **Retrieve:** In dieser Phase durchsucht der Benutzer die verschiedenen vom ihm ausgewählten Datenquellen.
- **Collate:** In dieser Phase erfolgt die strukturierte Ablage der gefundenen Informationen.
- **Interpret:** In dieser Phase nimmt der Benutzer die Interpretation der von ihm oder von Gruppenmitgliedern abgelegten Informationen vor.
- **Re-Present:** In dieser abschließenden Phase erfolgt die Darstellung des neu gewonnene Wissen, z.B. in Form einer wissenschaftlichen Publikation.

Die Umsetzung dieser fünf Phasen stellt ein Optimum dar und bietet den Benutzern ein in sich geschlossenes Serviceangebot. Strategische Unterstützung im Suchprozess erfährt der Benutzer in DAFFODIL durch eine Vielzahl von integrierten Werkzeugen, die sich durch höhere Suchfunktionen auszeichnen und das Dienstportfolio noch ergänzen. Auf einige diese Werkzeuge wird im Folgenden kurz eingegangen:

Referenzen und Zitationen

Das Referenzwerkzeug erlaubt ausgehend von einem bekannten Dokument, etwa aus der Resultatliste oder der persönlichen Handbibliothek, die Suche nach anderen Objekten, die von

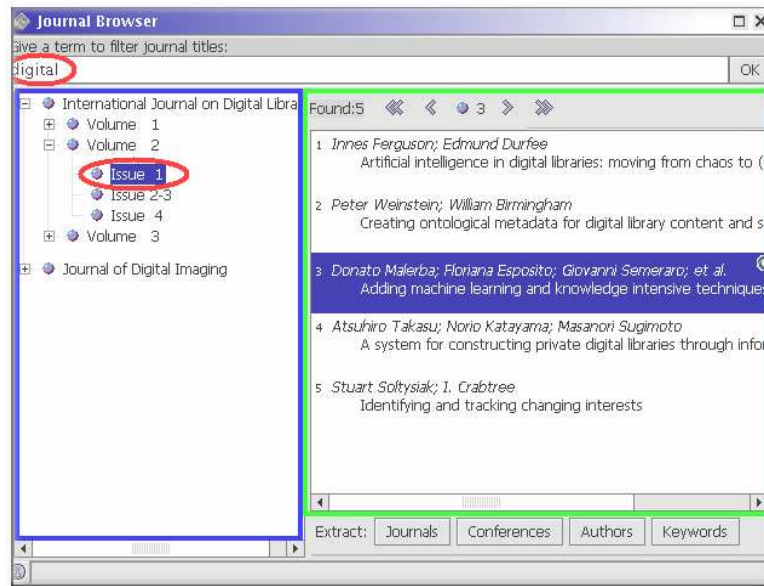


Abbildung 6: Journal-Werkzeug: Integriertes Suchen und Browsen in Journalen (Konferenzen)

Der Einsatz der Werkzeuge kann Ausgangspunkt oder Zwischenschritt eines umfangreicheren Suchplans sein. Explizite Verknüpfungen in den Detailansichten von Suchergebnissen weisen auf eine Zeitschrift oder einen Konferenzband hin, in dem ein Dokument veröffentlicht worden ist, und springen bei Aktivierung direkt in das Werkzeug.

Klassifikationen

Mit Hilfe des Klassifikationswerkzeuges erhalten Benutzer Zugriff auf eine hierarchische, themenorientierte Repräsentation des Suchraums. Es erlaubt das Browsen in Klassifikationsschemata wie dem ACM Computing Classification Scheme.

Thesauri

Über das Thesauruswerkzeug können zu Suchtermen generellere oder spezifischere Begriffe bzw. semantische Bedeutungen erfragt werden.

Fachspezifische oder webbasierte Thesauri werden für das Auffinden verwandter Begriffe genutzt. Die so gefundenen Begriffe können dann leicht in anderen Werkzeugen für weitere Anfragen benutzt werden.

Pro-aktivität und ihre Werkzeuge

Neben den bereits vorgestellten Werkzeugen im Bereich der Oberfläche und strategischen Unterstützung des Benutzers bei der Literatursuche innerhalb von DAFFODIL bietet das System auch eine Reihe an Werkzeugen, die pro-aktive Unterstützung bei der Suche, etwa bei der Formulierung von Anfragen anbieten. Diese sollen den Benutzer dabei etwa Hilfestellung geben, bestehende Anfrage zu verbessern, um mehr oder eindeutiger Ergebnisse bei einer Suche erzielen zu können.

Pro-aktive Anfrageunterstützung

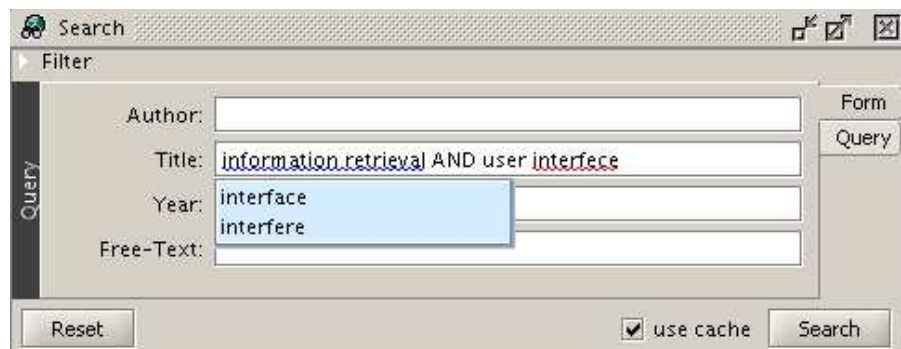
Die spezifischen Zielsetzungen der pro-aktiven Unterstützung bei der Anfrageformulierung sind:

1. Verringerung bzw. Vermeidung von grundlegenden Fehlern, syntaktischer und semantischer Art
2. Senkung der kognitiven Arbeitslast in frühen Phasen der Anfrageformulierungen
3. Erhöhung der Sicherheit bei der Anfragekonstruktion

Die Umsetzung der Ziele erhöht die Effizienz und Effektivität des Benutzers.

Als aktuelle Beispiele für die Umsetzung der angeführten Ziele im Bereich der pro-aktiven Anfrageunterstützung sollen hier die Rechtschreibkorrektur, die Korrektur semantischer Fehler und die Hilfestellungen z.B. im Bereich der verwandten Begriffe kurz dargestellt werden.

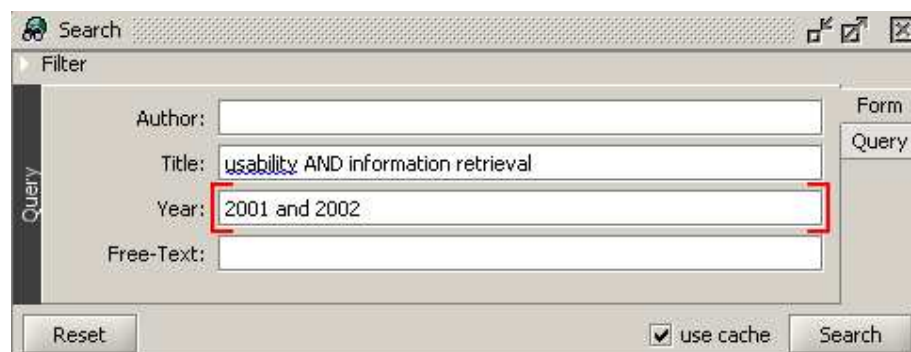
- *Rechtschreibkorrektur*: Eine Rechtschreibkontrolle macht Benutzer schon bei der Eingabe von Suchtermen auf möglicherweise falsch geschriebene Begriffe aufmerksam, indem es diese durch rote Unterschlägelung kenntlich macht. Als Hilfe bietet es auf Anfrage alternative oder korrigierte Schreibweisen des Begriffs an (s. Abbildung 7).



The screenshot shows a search window titled 'Search' with a 'Filter' dropdown. The 'Query' section contains four input fields: 'Author:', 'Title:', 'Year:', and 'Free-Text:'. The 'Title' field contains the text 'information retrieval AND user interfece', where 'interfece' is underlined in red. A blue tooltip box is positioned over the red underlined text, displaying the corrected word 'interfe' on two lines. To the right of the input fields are buttons for 'Form' and 'Query'. At the bottom of the window are buttons for 'Reset', a checked 'use cache' checkbox, and a 'Search' button.

Abbildung 7

- *Korrektur semantischer Fehler*: Häufige Denkfehler von Benutzern bei der Formulierung ihres Informationsbedürfnisses führen zu Anfragen, die objektiv semantische Fehler enthalten, und so keinen Ergebnissen führen können. Ein Beispiel für eine solche Anfrage ist in Abbildung 8 zu sehen. Hier wurde der Wunsch, Artikel aus den Jahren 2001 und 2002 zu finden, in die Boolesche Bedingung „2001 AND 2002“ übersetzt. Solche Fehler werden vom System bereits zur Eingabe erkannt und als problematisch markiert werden. Weitere Regeln semantischer Fehler sind leicht zu integrieren.



The screenshot shows a search window titled 'Search' with a 'Filter' dropdown. The 'Query' section contains four input fields: 'Author:', 'Title:', 'Year:', and 'Free-Text:'. The 'Title' field contains the text 'usability AND information retrieval'. The 'Year' field contains the text '2001 and 2002', which is enclosed in a red rectangular box. To the right of the input fields are buttons for 'Form' and 'Query'. At the bottom of the window are buttons for 'Reset', a checked 'use cache' checkbox, and a 'Search' button.

Abbildung 8

- *Begriffs-Unterstützung*: Als zusätzliche Hilfestellung bei der Formulierung und Reformulierung von Anfragen werden auch verwandte oder alternative Begriffe zu Suchtermen angezeigt (siehe Abbildung 9).

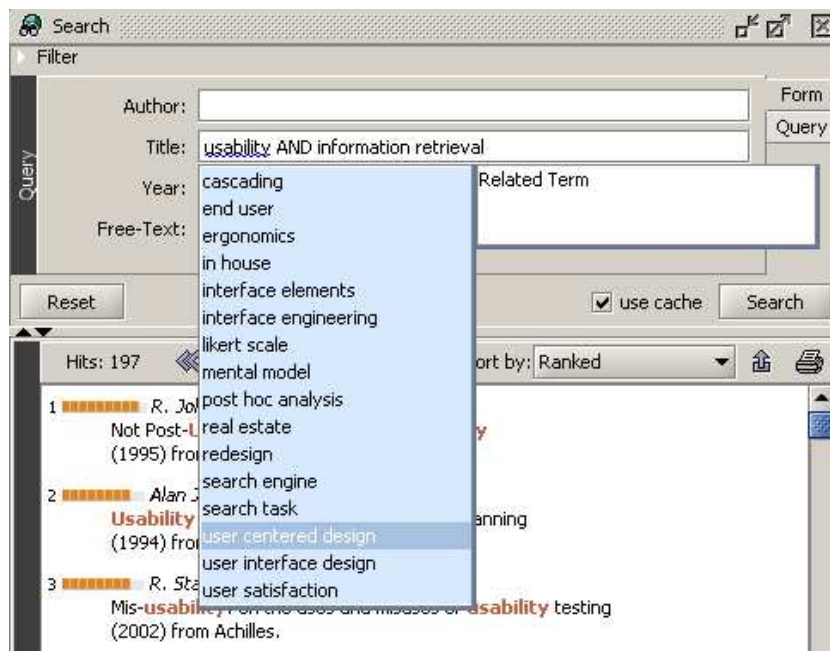


Abbildung 9

- *Begriffsvervollständigung*: Unter Nutzung früherer Suchanfragen und Eingaben des Benutzers können zusätzlich auch mögliche Vervollständigungen eines Teilbegriffes angeboten werden.

Pro-aktive Werkzeuge

Als zwei Beispiele für pro-aktive Werkzeuge im DAFFODIL-System können das Autorennamenwerkzeug und das 'Did You Mean?' Werkzeug genannt werden. Das erstere versucht bei einem unvollständig bekannten oder spezifizierten Namen (etwa mit fehlenden Vornamen) mögliche Ergänzungen vorzuschlagen. Das zweite Werkzeug schlägt Korrekturen für Anfragebegriffe, die über die Rechtsschreibkorrektur hinausgehen, vor.

Evaluationen

Im Rahmen der Projektlaufzeit sind zwei umfangreiche Evaluationen zur strategischen Unterstützung und zu den pro-aktiven Funktionen des DAFFODIL-Systems durchgeführt worden, um die Effektivitäts- und Effizienzsteigerungen zu belegen.

Strategische Unterstützung

Untersuchungsgegenstand bei der Evaluation der strategischen Unterstützung in Klas et al (2004) war die Frage, ob Benutzer mit DAFFODIL effizienter und effektiver den gestellten Aufgaben nachgehen können im Vergleich zu anderen Digitalen Bibliotheken. Die qualitative, mehrphasige Evaluation mit 28 Probanden hat das zugrunde liegende Modell und die daran geknüpften Erwartungen belegt. Durch die enge Integration höherer Suchfunktionalitäten mittels graphischer

Werkzeuge ermöglicht DAFFODIL effiziente und effektive Literaturrecherche und Informationsgewinnung in föderierten Digitalen Bibliotheken. Dabei hat sich abgezeichnet, dass die Effizienz- und Effektivitätssteigerung mit zunehmender Komplexität einhergeht.

Pro-aktive Unterstützung

Die Basis für die Evaluation der pro-aktiven Unterstützung stellen drei unterschiedlichen, qualitativen Benutzertests dar, dabei handelt es sich um:

- Eine heuristische Evaluation mit Folien von Beispielen für die Ansicht von pro-aktiven Funktionen.
- Einen kontrollierten Test mit zwei Studentengruppen mit und ohne pro-aktive Funktionen.
- Einen Test mit Einzelpersonen per Videoaufnahmen.

Die an die durchgeführten Tests gestellten Erwartungen umfassten das Verständnis, die Akzeptanz und den Einsatz der pro-aktiven Funktionen durch den Benutzer. Zudem sollten die Benutzer mit den pro-aktiven Funktionen erfolgreicher, d.h. effektiver und effizienter die gestellten Aufgaben bearbeiten.

Die Ergebnisse der heuristischen Evaluation zeigen, dass die Benutzer die pro-aktiven Funktionen akzeptiert und verstanden haben. Einige Aspekte des graphischen Designs der vorgestellten Beispiele haben einer anschließenden Überarbeitung bedurft. Interessant dabei erscheint die Forderung von Seiten der Benutzer in einem Abschlussgespräch nach weitergehenden und auch vollständig neuen pro-aktiven Funktionen.

Der nachfolgende Gruppentest hat mit 20 Informatikstudenten zwei Gruppen umfasst. Eine Gruppe hat von der Unterstützung durch pro-aktive Funktionen profitieren können, die andere Gruppe hat ähnliche Aufgaben, wie sie in Klas et al (2004) verwendet worden sind, ohne diese pro-aktiven Funktionen bearbeiten müssen.

Bei der Nutzergruppe ohne pro-aktive Funktionen hat sich gezeigt, dass ca. 45% der gestellten Anfragen fehlerhaft gewesen sind. Diese Fehler lassen sich in Rechtschreibfehler (Autorennamen und Anfrageterme), boolesche Anfragefehler, Übersetzungsprobleme und zu restriktive oder zu weitgefaste Anfragen differenzieren. Die erste Gruppe hat die Aufgaben dank der aktivierten pro-aktiven Funktionen mit deutlich mehr Sicherheit, weniger Nachfragen an die Besitzer und weniger negativen Kommentaren bearbeitet.

Der in der Abschlussphase dieser Evaluation durchgeführte Benutzertest ist mit der Videokamera dokumentiert worden, wodurch sich ein wesentlich tieferer Einblick in den kognitiven Prozess und in die Erwartungen der Benutzer ergeben hat als in den ersten beiden Tests möglich gewesen ist. Hierbei ist noch festzuhalten, dass das Verhalten der Benutzer erhebliche Variationen gezeigt hat, etwa bei der allgemeinen Akzeptanz der pro-aktiven Funktionen.

Zusammenfassend kann die Aussage gemacht werden, dass die pro-aktiven Funktionen positiv bewertet worden. Als wesentlicher Grund dafür ist die höhere Sicherheit der Benutzer bei der Anfrageerstellung zu nennen und die relativ schnelle Annahme des Komforts durch die Funktionen.

Resümee und Ausblick

Das Konzept der strategischen Unterstützung wird in DAFFODIL auf allen Ebenen des Systems realisiert; eine Handbibliothek bietet Personalisierung und Gruppenunterstützung. Ein Benutzer wird so in die Lage versetzt, komplexe Suchstrategien auf natürliche Weise zu verfolgen. Die Aufbereitung der Suchresultate durch DAFFODIL hilft bei der kognitiven Erfassung der Information. Pro-aktive Funktionen als eine weitere Ebene bieten umfangreiche Hilfestellungen an.

Dem Benutzer steht somit ein System zur Verfügung, das ihn in seinem Prozess der Literatursuche strategisch unterstützt, wodurch er gleichzeitig eine Steigerung der Effektivität und der Effizienz erfährt.

Unsere aktuellen Forschungsschwerpunkte umfassen die Bereiche der strategischen Unterstützung, der pro-aktiven Hilfe und der Awareness.

Innerhalb des Europäischen *Network of Excellence* DELOS wird DAFFODIL als Basisarchitektur für die Evaluation von Digitalen Bibliotheken genutzt. Des Weiteren wird an der Entwicklung eines einheitlichen System-Loggings gearbeitet, um qualitative Nutzerevaluationen durch quantitative Daten abzusichern. Dies ermöglicht darüber hinaus ein Vergleich von in Digitalen Bibliotheken verwendeten Werkzeugen und Diensten untereinander. DAFFODIL ist derzeit mit Erfolg im Institut für Informatik und Interaktive Systeme an der Universität Duisburg-Essen, Standort Duisburg, im Einsatz und wird von mehr als 50 Nutzern regelmäßig zur Literatursuche genutzt. Der Prototyp ist unter <http://www.daffodil.de> frei verfügbar.

Literatur

M. J. Bates. Information search tactics. *Journal of the American Society for Information Science*, 30(4):205-214, 1979.

Marcia J. Bates. Where should the person stop and the information search interface start? *Information Processing and Management*, 26(5):575-591, 1990.

N. Fuhr, N. Gövert, and C.-P. Klas. An agent-based architecture for supporting high-level search activities in federated digital libraries. In *Proceedings 3rd International Conference of Asian Digital Library*, pages 247-254, Taejon, Korea, 2000. KAIST.

Norbert Fuhr, Claus-Peter Klas, Andre Schaefer, and Peter Mutschke. Daffodil: An integrated desktop for supporting high-level search activities in federated digital libraries. In *Research and Advanced Technology for Digital Libraries. 6th European Conference, ECDL 2002*, pages 597-612, Heidelberg et al., 2002. Springer.

Claus-Peter Klas, Norbert Fuhr, and Andre Schaefer. Evaluating strategic support for information access in the DAFFODIL system. In Rachel Heery and Liz Lyon, editors, *Research and Advanced Technology for Digital Libraries. Proc. European Conference on Digital Libraries (ECDL 2004)*, Lecture Notes in Computer Science, Heidelberg et al., 2004. Springer.

A. Paepcke. Digital libraries: Searching is not enough - what we learned on-site. *D-Lib Magazine*, 2(5), May 1996. <http://www.dlib.org/dlib/may96/stanford/05paepcke.html>.

S. Weibel and E. Miller. A summary of the CNI/OCLC image metadata workshop. *D-Lib Magazine*, 3(1), January 1997. <http://www.dlib.org/dlib/january97/oclc/01weibel.html>