



Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Adaptive Retrieval, Composition & Presentation of Closed-Corpus and Open-Corpus Information

A thesis submitted to the

University of Dublin, Trinity College

for the degree of

Doctor of Philosophy

Ben Steichen

Knowledge and Data Engineering Group

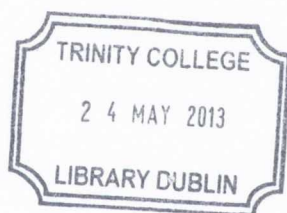
Intelligent Systems

School of Computer Science and Statistics

Trinity College Dublin

Ireland

2012



Thesis 10020

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

Acknowledgements

I would like to deeply thank my supervisor Prof. Vincent Wade, whose expert knowledge, guidance and dedication has mentored me throughout this work. I would also like to express my gratitude to Prof. Helen Ashman for her insightful feedback.

A huge thank you to all my colleagues and friends in the Knowledge and Data Engineering Group (KDEG), many of whom I had the chance to collaborate with over the course of this research. I would also like to extend my appreciation to all members of the Centre for Next Generation Localisation (CNGL). In particular, I would like to express my gratitude to Fred Hollowood, Johann Roturier and Jason Rickard from Symantec for all their contributions and help during this research.

Most importantly, I would like to thank Bo and my parents for their unconditional belief, support and encouragement throughout the years.

Abstract

A key challenge for information access systems lies in their ability to deliver information that is most suited to a user's needs, preferences and context. Personalised Information Retrieval (PIR) seeks to address this challenge by tailoring the selection of results to each individual user. Such PIR systems typically generate adaptive result rankings based on historic user interests or location properties. However, other considerations such as user needs, preferences or context are often neglected. Moreover, users are typically only presented with linear (monolingual) result rankings that do not provide any adaptive navigation support across different information sources. On the other hand, the field of Adaptive Hypermedia (AH) has inherently focused on generating non-linear, hyperlinked result compositions. This enables adaptive navigation and presentation support, allowing users a guided experience through an information space. Moreover, AH systems typically generate adaptive responses according to multiple considerations (also called personalisation "dimensions"), such as user needs, knowledge and context. However, AH techniques have typically only been applied across closed-corpus content bases, requiring substantial amounts of metadata. The key problem remains in providing such adaptive compositions across open-corpus information sources (in addition to closed corpora).

In order to address this problem, the thesis presents a novel compositional approach to open- and closed-corpus information retrieval and delivery through an innovative combination of Adaptive Hypermedia and Personalised Information Retrieval techniques. This technology enables the first dynamic integration and multidimensional adaptation of multilingual open and closed corpora. In particular, the contribution of the thesis is an extension of PIR and AH techniques to enable informed multiple adaptive query generation and adaptive result recomposition and presentation. This innovation is evaluated and validated through a series of case study implementations and evaluations, which show that the compositional approach successfully supports authentic user information needs in a personalised manner. In particular, it is shown that users are more efficient, effective and satisfied with the compositional approach compared to conventional information retrieval systems. Moreover, the approach is shown to be able to support multiple dimensions of adaptation, including user intent, language, knowledge, interface preferences and device capabilities.

Table of Contents

Declaration.....	i
Acknowledgements.....	ii
Abstract.....	iii
Table of Contents.....	iv
Table of Figures.....	vii
Table of Tables.....	ix
1 Introduction.....	1
1.1. Motivation.....	1
1.2. Research Question.....	4
1.3. Objectives.....	5
1.4. Methodology.....	5
1.5. Contribution.....	6
1.6. Thesis Overview.....	8
2 Adaptive Hypermedia & Personalised Information Retrieval.....	10
2.1. Introduction.....	10
2.2. Query adaptation.....	13
2.2.1. Summary and Critique.....	18
2.2.2. Comparison across Query Adaptation techniques.....	21
2.3. Retrieval adaptation.....	23
2.3.1. Statistical methods.....	24
2.3.2. Metadata-based approaches.....	26
2.3.3. Summary and Critique.....	30
2.3.4. Comparison across Retrieval Adaptation techniques.....	32
2.4. Adaptive Composition & Presentation.....	35
2.4.1. Statistical techniques.....	35
2.4.2. Metadata-based techniques.....	38
2.4.3. Summary and Critique.....	43
2.4.4. Comparison across Adaptive Composition & Presentation techniques....	44
2.5. Conclusions.....	47
2.5.1. User dimensions.....	47
2.5.2. Adaptation techniques.....	49
2.5.3. Overall Findings and Complementary Affordances.....	51
3 Initial Adaptive Open-Corpus Composition System.....	54
3.1. Introduction.....	54
3.2. Contribution of the author.....	55
3.3. Architecture.....	56
3.3.1. Models.....	56
3.3.2. Architecture Components & Capabilities.....	57
3.3.3. Technological Architecture.....	59
3.4. Prototype Implementation.....	61
3.4.1. Prototype Prerequisites.....	61
3.4.2. Adaptation Process.....	64
3.5. Evaluation.....	68
3.5.1. Educational Benefit and User Satisfaction Hypotheses.....	68
3.5.2. Comparison to baselines.....	69
3.5.3. Experimental Setup.....	71
3.5.4. Results.....	75
3.5.5. Discussion.....	80
3.6. Conclusions.....	81

4	ARCHING - Adaptive Retrieval and Composition of Heterogeneous INformation sources for personalised hypertext Generation.....	84
4.1.	Introduction	84
4.2.	Design Principles.....	85
4.3.	Architecture	87
4.3.1.	Architecture Components & Capabilities.....	87
4.3.1.1.	Search	88
4.3.1.2.	Expand.....	88
4.3.1.3.	Translate	89
4.3.1.4.	Classify.....	89
4.3.1.5.	Additional Design Flexibility	90
4.3.2.	Technological Architecture	90
4.3.2.1.	Search	91
4.3.2.2.	Expand.....	94
4.3.2.3.	Translate	95
4.3.2.4.	Classify.....	95
4.3.2.5.	Additional Design flexibility	96
4.4.	Prototype Implementation	98
4.4.1.	Personalised Customer Care.....	98
4.4.2.	Prototype Ecosystem	99
4.4.3.	Adaptation Processes.....	102
4.4.3.1.	Adaptation Process Models	103
4.4.3.2.	Overview Results - Adaptation Process	104
4.4.3.3.	Detailed Results - Adaptation Process.....	108
4.5.	Evaluation.....	112
4.5.1.	Hypotheses/Sub hypotheses	112
4.5.2.	Comparison to Baseline.....	114
4.5.3.	Experimental Setup.....	115
4.5.4.	Results	117
4.5.5.	Discussion.....	124
4.6.	Conclusions	126
5	Investigation & Evaluation of Open-Web Personalisation: A Comparative Approach	129
5.1.	Introduction	129
5.2.	Open-Web Information Source Selection.....	130
5.3.	Design.....	132
5.3.1.	Design of Interface Composition 1: Information Source Panels.....	133
5.3.2.	Design of Interface Composition 2: Topic-based Composition (Highly structured).....	135
5.3.3.	Design of Interface Composition 3: Topic-based Composition (Loosely structured).....	138
5.3.4.	Discussion.....	140
5.4.	Implementing the interface compositions.....	141
5.4.1.	Interface Composition 1 Implementation (C1).....	142
5.4.2.	Interface Composition 2 Implementation (C2).....	144
5.4.3.	Interface Composition 3 Implementation (C3).....	145
5.5.	Evaluation of interface compositions	146
5.5.1.	Hypotheses/Sub hypotheses	147
5.5.2.	Experimental Setup.....	149
5.5.3.	Results	151
5.5.4.	Overall Evaluation Findings.....	164
5.6.	Conclusions	166

6	Multilingual Information Composition & Additional Adaptation Dimensions.....	168
6.1.	Introduction.....	168
6.2.	Multilingual Information Access	169
6.3.	Personalised Multilingual Customer Care	170
6.4.	Design and implementation of Multilingual Composition Prototypes	171
6.4.1.	Multilingual Composition 1 (MC1).....	171
6.4.2.	Multilingual Composition 2 (MC2).....	172
6.4.3.	Multilingual Composition 3 (MC3).....	174
6.5.	Evaluation of Multilingual Composition Prototypes	175
6.5.1.	Hypotheses/Sub hypotheses.....	175
6.5.2.	Experimental Setup.....	177
6.5.3.	Results.....	178
6.5.4.	Overall Evaluation Findings & Additional Support Possibilities	189
6.6.	Additional Adaptation Dimensions and Prototypes.....	193
6.6.1.	Mobile device adaptation.....	193
6.6.2.	Fine-grained user expertise modelling.....	194
6.6.3.	Multimedia support.....	195
6.6.4.	Alternative customer care application.....	196
6.7.	Conclusions.....	198
7	Conclusions.....	199
7.1.	Introduction.....	199
7.2.	Research question and objectives	199
7.3.	Contributions.....	205
7.4.	Future Work	207
	Bibliography.....	211
	APPENDIX A. SQL Experiment (see section 3.5).....	A-1
	A.1. Task Questions.....	A-1
	A.2. System Usability Scale (SUS) (Brooke, 1996).....	A-2
	A.3. Additional Questions.....	A-3
	APPENDIX B. Symantec Experiment 1 (see section 4.5).....	B-1
	B.1. Task Questions.....	B-1
	B.2. User Satisfaction regarding system-specific features	B-4
	B.3. General User Satisfaction, Reaction & Comments	B-5
	B.4. Comparative Questionnaire 1.....	B-6
	B.5. Comparative Questionnaire 2.....	B-7
	APPENDIX C. General Support Questions (see section 5.2).....	C-1
	APPENDIX D. Product reaction cards (see section 5.3).....	D-1
	APPENDIX E. Symantec Experiment 2 (see section 5.5).....	E-1
	E.1. Task questions.....	E-1
	E.2. Pre-Questionnaire.....	E-2
	E.3. Interface Preference Prequestionnaire (asked for each of the screens).....	E-3
	E.4. Post-task Questionnaire.....	E-6
	E.5. Application-specific Questionnaire.....	E-7
	APPENDIX F. Full Pre-questionnaire results (see section 5.5.3)	F-1
	APPENDIX G. Symantec Experiment 3 (see section 6.5)	G-1
	G.1. Task questions.....	G-1
	G.2. Pre-Questionnaire.....	G-2
	G.3. Post-task Questionnaire.....	G-4
	G.4. Comparative Questionnaire.....	G-5

Table of Figures

Figure 2-1. User profile generation using ODP categories	15
Figure 2-2. User interface for manual selection of additional query terms (Ahn et al., 2010).....	16
Figure 2-3. Personalised meta-search engine (Glover et al., 2001).....	18
Figure 2-4. Example of a faceted search interface (Yee, et al., 2003)	39
Figure 2-5. Example of link annotation (Hsiao et. al., 2009)	42
Figure 3-1. Adaptive Engine Component Architecture.....	58
Figure 3-2. Adaptive Engine Technological Architecture.....	60
Figure 3-3. Content Harvesting and Metadata Generation process.....	62
Figure 3-4. SQL ontology.....	63
Figure 3-5. Query Elicitation.....	64
Figure 3-6. Composition generation process.....	65
Figure 3-7. Result overview screen.....	67
Figure 3-8. Content introducing the SQL Privilege concept.....	67
Figure 3-9. IR+RF prototype: top of result screen	70
Figure 3-10. IR+RF prototype: bottom of screen, with query refinement button	70
Figure 3-11. IR+AH prototype architecture	71
Figure 3-12. Pre-questionnaire	73
Figure 3-13. Pre-task questions	74
Figure 3-14. Task screen	75
Figure 3-15. Duration of annotation events that produced complete descriptions in under five minutes	76
Figure 3-16. Number of documents that users looked at per query	77
Figure 4-1. ARCHING Component Architecture.....	87
Figure 4-2. ARCHING Technological Architecture	91
Figure 4-3. Asynchronous Open Corpus Utility calls	97
Figure 4-4. Part of the DocBook structure modelled as ontology classes.....	100
Figure 4-5. Domain ontology of product features	101
Figure 4-6. Query Elicitation.....	103
Figure 4-7. Result overview generation process	104
Figure 4-8. Closed-Corpus Retrieval and Metadata Retrieval	105
Figure 4-9. Information Grouping & Open-Corpus Query Generation.....	106
Figure 4-10. Result Overview Screen.....	106
Figure 4-11. Overview Results for “Backup”, including related features (top).....	107
Figure 4-12. Result Model Transformation & Asynchronous Open-Corpus Retrieval.....	107
Figure 4-13. Detailed Results - Adaptation Process.....	108
Figure 4-14. Information Grouping & Additional Result Retrieval.....	109
Figure 4-15. Detailed Results for the “Backup” feature, currently displaying knowledge item from documentation.....	110
Figure 4-16. Structured Navigation.....	110
Figure 4-17. Result Model Transformation & Asynchronous Open-Corpus Retrieval.....	111
Figure 4-18. Detailed Results for the “Backup” feature, currently displaying support forums.....	112
Figure 4-19. Non-adaptive search system	115
Figure 4-20. Experimental Process.....	117
Figure 4-21. Task Completion times	118
Figure 4-22. Number of queries	119
Figure 4-23. Information viewed.....	120
Figure 4-24. User satisfaction.....	122

Figure 4-25. Comparative user perception of result presentation, composition and personalisation	122
Figure 4-26. User motivation, engagement and fun.....	124
Figure 5-1. Composition 1 Mockup.....	133
Figure 5-2. Composition 2a Mockup.....	135
Figure 5-3. Composition 2b Mockup.....	136
Figure 5-4. Composition 3a Mockup.....	138
Figure 5-5. Composition 3b Mockup.....	139
Figure 5-6. Composition 1 Adaptation Process	142
Figure 5-7. Composition 1 Screenshot.....	143
Figure 5-8. Composition 2 Screenshot (second screen).....	145
Figure 5-9. Composition 3 Screenshot (second screen).....	146
Figure 5-10. Experimental process	151
Figure 5-11. Number of Queries.....	154
Figure 5-12. Information Viewed	155
Figure 5-13. Perceived Task Complexity	156
Figure 5-14. Perceived Success	157
Figure 5-15. Perceived Task Assistance	158
Figure 6-1. Multilingual Composition 1 (MC1) Screenshot.....	172
Figure 6-2. Multilingual Composition 2 (MC2) Screenshot (overview screen).....	173
Figure 6-3. Multilingual Composition 2 (MC2) Screenshot (detailed screen)	173
Figure 6-4. Multilingual Composition 3 (MC3) Screenshot.....	174
Figure 6-5. Experimental Process	178
Figure 6-6. Time on task.....	180
Figure 6-7. Number of queries.....	181
Figure 6-8. Information Viewed	182
Figure 6-9. Task Success	183
Figure 6-10. Perceived Task Assistance	184
Figure 6-11. “Quick-Fix” vs “Understanding a problem”	186
Figure 6-12. “Sequential” vs. “global” learning	186
Figure 6-13. Understanding through individual parts vs holistic view.....	186
Figure 6-14. User Satisfaction regarding multilingual information composition.....	188
Figure 6-15. Direct comparison between multilingual compositions.....	188
Figure 6-16. General Multilingual Information Access Preferences	189
Figure 6-17. Multilingual Composition with translation support	192
Figure 6-18. Multilingual Composition with French, English and German results.....	192
Figure 6-19. Mobile device compositions.....	194
Figure 6-20. User expertise modelling.....	195
Figure 6-21. Information composition including video results.....	196
Figure 6-22. Microsoft Office prototype.....	197

Table of Tables

Table 2-1. Summary of Query Adaptation Techniques.....	21
Table 2-2. Summary of Retrieval Adaptation Techniques	33
Table 2-3. Summary of Adaptive Composition and Presentation Techniques.....	45
Table 5-1. Information Source Preferences.....	132
Table 5-2. User reaction cards for Interface Composition 1 Design Mockup.....	134
Table 5-3. User reaction cards for Interface Composition 2 Design Mockup.....	137
Table 5-4. User reaction cards for Interface Composition 3 Design Mockup.....	140
Table 5-5. Prequestionnaire website screenshots, used to capture general search interface preferences.....	150
Table 5-6. Overall prequestionnaire answers	152
Table 5-7. Overall task characteristics	153
Table 5-8. Interface Preference Correlations.....	159
Table 5-9. Additional characteristic correlations	160
Table 5-10. Perceived Navigation, Composition and Presentation	161
Table 5-11. Perceived usage and appropriateness of query intent	162
Table 5-12. User motivation and frustration	163
Table 6-1. Overall prequestionnaire answers for the multilingual study	179
Table 6-2. Perceived Task Assistance	184
Table 6-3. User Satisfaction regarding multilingual information composition.....	187

1 Introduction

1.1. Motivation

A key challenge for information access systems lies in their ability to deliver information that is most suited to a user's needs, preferences and context (Brusilovsky, et al., 2007). The concepts of adaptation and personalisation are increasingly emerging on today's web in order to address this challenge. Obvious examples of personalisation can be found in modern e-commerce systems such as Amazon¹, which recommend items to users based on prior viewing and purchase behaviour. More covert types of personalisation can be observed in modern search engines such as Google², which alter result rankings based on a user's search history and clickthrough behaviour.

There are both benefits and dangers in applying personalisation in information access systems. Clear benefits have been reported in domains such as e-learning, where users receive personalised guidance through learning material (Conlan and Wade, 2002) (Brusilovsky, et al., 2004) (De Bra, et al., 2003). Such benefits to the user can range from increased learning effectiveness to improved user satisfaction (Conlan and Wade, 2004), as well as increased user motivation (Hsiao et al., 2009). On the other hand, the dangers of personalisation can range from security and data privacy concerns (Ashman, et al., 2009), as well as to the so-called "filter bubble" (Pariser, 2011), whereby users only receive personalised information streams and fail to get access to contrasting opinions and viewpoints (Billsus and Pazzani, 2007).

The need for adaptation in search systems has been identified by many researchers in the field of Information Retrieval. In (Jansen et al., 2007), web queries are classified into three categories, *Navigational searching*, *Transactional searching* and

¹ <http://www.amazon.com>

² <http://www.google.com>

Informational searching. Navigational queries typically only require one precise answer, for example the web address of a particular company or the homepage of an individual. Similarly, transactional queries are very specific, as the main goal is to purchase a particular product or use a particular service. These two types of queries are shown to constitute only 20% of current searches on the web (Jansen et al., 2008). The remaining 80% of queries can be classified as Informational Searching, i.e. a user searching for comprehensive information on a particular topic. Current IR solutions, although being successful in improving the accuracy of ranked result lists, do not assist this Informational searching adequately, as often more than a few very precise results are needed to fill a user's knowledge gap. More adaptive solutions are required, which present users with comprehensive and personalised information compositions in order to provide a guided experience through an information space.

White and Morris (2007) investigate the differences in behaviour between search experts and novices, particularly the use of advanced search features (e.g. AND, OR, ""). They found several differences between the two groups, indicating that search expertise greatly influences search behaviour. Similarly, in (White, et al., 2009) a large query log is analysed in order to find more differences between expert users and non-expert users. Again, there are noticeable differences in search behaviour, such as experts visiting less commercial sites or using more domain-specific terms. Wildemuth (2004) analyses differences in search behaviour for students before and after they have acquired domain knowledge. Common patterns are shown to exist for both cases, such as adding and deleting terms to search queries. However, after having acquired more domain knowledge, students were generally quicker at selecting the right terms. Additionally, with increasing domain knowledge, the number of refinement moves was observed to be lower.

The above studies provide clear evidence that the one-size-fits-all paradigm of traditional Information Retrieval systems does not address the various differences in user information needs, preferences and context³.

The application of Personalised Information Retrieval (PIR) that can be found in modern search engines attempts to overcome some of these shortcomings. Such

³ Context, as defined in (Dey, 2001) "is any information that can be used to characterise the situation of an entity." While such information can include data from physical sensors such as GPS or accelerometers (e.g. in Brown and Jones (2001), Maekawa et al. (2009), Noh, et al. (2011)), the notion of context within this thesis relates to the situational information need of a person, e.g. including the current task or the person's domain knowledge.

systems typically represent users with simplified personas, which are often based on historic interests or user location properties (e.g. geographical location, language prevalent in a region). Using such data, statistical approaches for query adaptation and result reranking enable the efficient personalisation of search results. However, other considerations such as user preferences or context are often neglected. Moreover, users are typically only presented with linear result rankings that do not provide any adaptive navigation support across the various information sources to satisfy an informational query.

By moving towards adaptation and personalisation, PIR increasingly faces challenges that have been addressed extensively by the field of Adaptive Hypermedia (AH). The approaches and techniques that have arisen in AH have been inherently focused on generating personalised responses according to varying user needs, backgrounds and contexts. Such an integration of multiple considerations, also called personalisation “dimensions” (Wade, 2009), allows the non-linear composition of results according to a user’s current goals, preferences and context. AH techniques can therefore be used to produce adaptive composition, navigation and presentation support, allowing users a more guided experience through an information space. This non-linear approach of AH with respect to composition stems from its background in hypertext research, as opposed to the document-centric ranking paradigm of traditional IR systems.

However, there have been substantial shortcomings to AH techniques as well, most notably in terms of applicability to open-corpus information. Traditionally, AH techniques apply their adaptation over a closed-corpus content base, requiring substantial amounts of metadata. Open Adaptive Hypermedia (OAH) research tries to address this limitation by producing adaptive compositions from open-corpus information sources (Brusilovsky and Henze, 2007). Current OAH systems attempt such compositions using either manual, collaborative or automatic methods. Manual (authoring) approaches typically allow a user to identify and include open resources at design time. The problem with such approaches is that leveraging and annotating such information requires significant manual effort and all the information needs to be generated *a priori*. Collaborative techniques provide guidance across open-corpus information by deriving the selection of relevant information from the quantity of users stepping between content. However, the gap with such systems remains in terms of maintaining user guidance across the conceptual domain through coherent adaptive navigation techniques. Automatic linking techniques partially overcome the scalability

limitations by automatically estimating the relatedness between pages. However, such techniques have typically taken a form similar to recommender systems, failing to fully combine the various information sources (closed-corpus and open-corpus) into an integrated adaptive composition.

Another gap of AH learning systems to date is that they typically focus on producing educational course compositions based on a predefined (educational) need, rather than addressing an informal need indicated by a user search query. More dynamic (ad-hoc) user needs, which represent more closely the type of information requests currently found on the web, are typically not supported.

A key problem hence remains in satisfying an informal user need through an adaptive information composition from closed-corpus and open-corpus information sources, adapted according to multiple user dimensions. In order to generate such information compositions, next generation information access systems need to not only perform query adaptation and result reranking of open-corpus information, but also to recompose closed-corpus and open-corpus information into an integrated adaptive presentation according to multiple user dimensions.

1.2. Research Question

This thesis is researching the techniques and technologies required to generate adaptive information compositions that satisfy informal queries according to multiple dimensions⁴ of adaptation across closed-corpus and open-corpus information sources. More specifically, it asks the question, **“what adaptive techniques and technologies are needed to provide such multidimensional information compositions across closed and open corpora in order to enhance a user’s *effectiveness, efficiency and satisfaction*⁵.”**

In this thesis, a closed corpus refers to the general Adaptive Hypermedia definition of a “closed corpus of documents, where documents and relationships between the documents are known to the system at design time” (Brusilovsky and Henze, 2007). We

⁴ Examples of such adaptation dimensions include user intent, language preferences, prior knowledge, device capabilities, etc. (Wade, 2009)

⁵ The benefits to the user are defined in this thesis according to the three aspects of usability identified by the International Standards Organization [ISO 9241-11] as “the extent to which a product can be used by specified users to achieve specified goals with *effectiveness, efficiency and satisfaction* in a specified context of use.”

extend this definition in this thesis as “a corpus that consists of highly structured data and contains multiple levels of conceptual models and metadata”. Open-corpus generally refers to “a set of documents that is not known at design time and, moreover, can constantly change and expand” (Brusilovsky and Henze, 2007). We extend this notion in this thesis and define it as “any information available on the open web, including corporate websites, social media or other forms of content such as videos and images”.

1.3. Objectives

In order to address the research question discussed in section 1.2, the following objectives have been identified.

- Identify key affordances, techniques and impacts of current adaptive information access systems, particularly in the areas of Adaptive Hypermedia (AH) and Personalised Information Retrieval (PIR).
- Design and develop system architectures and adaptation processes that enable the generation of adaptive information compositions according to multiple levels of adaptation across closed-corpus and open-corpus information sources.
- Evaluate the architectures through a series of case-study implementations using metrics related to *user efficiency*, *effectiveness* and *satisfaction*.

1.4. Methodology

As stated in sections 1.1 and 1.2, this research focuses on techniques and technologies required to generate adaptive information compositions across closed-corpus and open-corpus information sources. Furthermore, such compositions need to be able to satisfy informal queries according to multiple dimensions of adaptation.

First of all, the stated requirements demand a thorough investigation of techniques and technologies in personalised information access systems, in particular Personalised Information Retrieval, as well as Adaptive Hypermedia. A state of the art review of techniques and technologies from these fields is presented in chapter 2.

Following this investigation, the methodology applied in this research consists of an iterative case-study-based approach, whereby a series of adaptive information composition architectures and prototypes are designed and developed. Each cycle includes a thorough evaluation of the research prototypes in authentic use case scenarios and measures the benefits to users in terms of user efficiency, effectiveness and satisfaction. The analysis of the results in each evaluation cycle then enables further refinements to the developed architectures and prototypes.

The evaluation methodology chosen in this research consists of task-based user experiments, whereby the research prototypes are assessed through comparative evaluations. The particular method of task-based user evaluation has been chosen because it represents an effective method for measuring the impact of techniques and technologies on user performance in realistic real-world scenarios (He et al. 2008). Although lab-based precision and recall measurements still dominate the Information Retrieval research field, many research works have called for more user-centric measures, such as task-based evaluation methods (He et al. 2008) or interactive information retrieval evaluation models (Borlund 2003). It is argued that realistic scenarios cannot be supported by traditional batch evaluations and that the real system performance can only be measured by placing the user at the centre of the evaluation (Borlund 2003). Since this thesis aims to research particularly the benefits to the user, a task-based user evaluation methodology constitutes an effective method of assessing the prototype performances.

1.5. Contribution

The major scientific contribution of this thesis is a novel *compositional approach to open- and closed-corpus mono- and multilingual* information retrieval and delivery through a combination of Adaptive Hypermedia and Personalised Information Retrieval techniques. This approach enables the first application technology that can support the dynamic integration and multidimensional adaptation of multilingual open and closed corpora. In particular, the novel innovation is an extension of PIR and AH techniques to enable informed multiple adaptive query generation and adaptive result recomposition and presentation. The approach thereby enables the application of adaptive navigation across closed-corpus and open-corpus information according to multiple dimensions of adaptation.

The minor contribution consists of the demonstration of this innovation through a series of case study implementations and evaluations, which show that the compositional approach successfully supports authentic user information needs in a personalised manner. In particular, it is shown that users are more efficient, effective and satisfied with the compositional approach compared to conventional information retrieval systems. Moreover, the approach is able to support multiple dimensions of adaptation, including user intent, language, knowledge, interface preferences and device capabilities.

In order to evaluate and validate the benefits of the compositional approach across heterogeneous information sources, a generic architecture called ARCHING (Adaptive Retrieval and Composition from Heterogeneous Information for personalised hypertext Generation) has been developed, which fully retains AH capabilities, while integrating adaptive open-corpus manipulation capabilities. Through a number of prototype implementations and evaluations, it is shown that this architecture can be used to successfully implement different types of information composition prototypes in order to suit particular user preferences and characteristics.

These contributions have resulted in a number of high-quality conference publications:

Steichen, B., O'Connor, A., and Wade, V. (2011). Personalisation in the Wild - Providing Personalisation across Semantic, Social and Open-Web Resources. *In Proceedings of the 22nd ACM conference on Hypertext and hypermedia (HT '11)*, pp. 73-82.

Steichen, B., and Wade, V. Adaptive Retrieval and Composition of Socio-Semantic Content for Personalised Customer Care. (2010). *In Proceedings of the International Workshop on Adaptation in Social and Semantic Web (SAS-WEB 2010), in conjunction with UMAP 2010, CEUR Workshop Proceedings, ISSN 1613-0073*, pp. 1-10.

Steichen, B., Lawless, S., O'Connor, A., and Wade, V. (2009). Dynamic hypertext generation for reusing open corpus content. *In Proceedings of the 20th ACM conference on Hypertext and hypermedia (HT '09)*, pp. 119-128.

The second minor contribution consists of a novel comparison of Personalised Information Retrieval and Adaptive Hypermedia, which i) surveys the key techniques

and technologies ii) analyses their respective strengths and weaknesses and iii) identifies the potential for integrating complementary affordances. This contribution has resulted in a high-quality journal publication:

Steichen, B., Ashman, H., and Wade, V. (2012). A comparative survey of Personalised Information Retrieval and Adaptive Hypermedia techniques. *Information Processing and Management*, Vol. 48, Issue 4, pp. 698–724.

1.6. Thesis Overview

The remainder of this thesis is organised as follows. Chapter 2 identifies and discusses the key affordances, techniques and impacts in the use of Adaptive Hypermedia and Personalised Information Retrieval technologies. In particular, techniques and technologies are discussed across three search process adaptation stages, namely query adaptation (section 2.2), adaptive retrieval (section 2.3) and adaptive composition and presentation (section 2.4).

Chapter 3 describes the first iteration of an adaptive open-corpus composition system using a state-of-the-art Adaptive Hypermedia architecture. The evaluation of an implemented prototype in an authentic eLearning scenario shows that such a system can be applied successfully for solving learning tasks using externally-sourced open-corpus data. It is shown that the system motivates users to explore more resources while issuing the same number of queries as with standard search systems. However, several shortcomings of the architecture are also highlighted, most notably its strong reliance on metadata for retrieving and composing information. In particular, the integrated open-corpus information needs to be fully marked up *a priori* and requires significant manual effort. This retrieval limitation also restricts the user during the query elicitation stage, as only concepts known to the system can be used as query inputs. Therefore, in order to satisfy informal keyword queries and to integrate open-corpus information on-the-fly, it is argued that the architecture needs to be extended with more lightweight open-corpus retrieval capabilities.

Motivated by the background findings in chapter 2 and the conclusions drawn from the experimental findings in chapter 3, chapter 4 presents an architecture for Adaptive Retrieval and Composition of Heterogeneous Information for personalised hypertext Generation (ARCHING). This architecture fully retains the Adaptive Hypermedia

capabilities of the first iteration, while integrating open-corpus retrieval and adaptation capabilities. This architecture allows a novel compositional approach to information retrieval and delivery, which can generate adaptive information compositions from closed-corpus and open-corpus information sources according to multiple dimensions of adaptation. An evaluation in an authentic customer care scenario is described, which compares a prototype implementation of ARCHING to a non-adaptive, purpose-built search system. Results from this task-based evaluation show that the prototype significantly enhances a user's efficiency, effectiveness and satisfaction.

Motivated by the positive findings from this evaluation, chapter 5 presents a series of distinct interface compositions, each implemented using the ARCHING architecture. Moreover, these compositions each integrate information that is retrieved and composed on-the-fly from the open web. Comparative evaluations are carried out in order to determine the efficiency, effectiveness and user satisfaction regarding each of the prototypes. The evaluations analyse varying user characteristics and show that different interface compositions suit different users and task contexts.

In addition to the adaptation possibilities presented in chapters 4 and 5, chapter 6 presents a multitude of further adaptation dimensions that can be supported using the compositional approach. In particular, this chapter first investigates the degree to which the compositional approach can be used to support the dimension of user language competencies through multilingual information compositions. Secondly, the chapter demonstrates a number of additional dimensions supported by ARCHING, such as different device interfaces, user expertise modelling and multimedia preferences.

Finally, chapter 7 concludes this thesis with a summary of the findings and contributions of this research. Moreover, a number of future research directions in adaptive information retrieval, composition and presentation are suggested.

2 Adaptive Hypermedia & Personalised Information Retrieval

2.1. Introduction

This chapter analyses the key affordances, techniques and impacts in the use of Adaptive Hypermedia (AH) and Personalised Information Retrieval (PIR). The comparison identifies their respective strengths and weaknesses, as well as the potential for the fusion of selected techniques and approaches. This fusion of techniques aims to enable the affordances outlined in section 1.2. “to generate adaptive information compositions that satisfy informal queries according to multiple dimensions of adaptation across closed-corpus and open-corpus information sources.” By contrast, current PIR systems typically apply personalisation over large open-corpus data on the sole dimension of previous interests, whereas AH systems typically apply multi-dimensional personalisation only over the closed-corpus data that they manage.

Both research fields have traditionally aimed to solve similar issues and challenges but approached from opposite directions: PIR provides personalised information predominantly through adaptive document ranking techniques, whereas AH delivers personalised information through adaptive compositions and presentations. This difference between AH and PIR stems from the overall information access paradigms underlying both areas. Moreover, this distinction has driven most of the research in both fields and has led to their respective techniques and technologies towards adaptation and personalisation.

AH systems stem from the information access paradigm of searching by browsing, where users generally have less precise information needs and therefore need to browse and explore pages. AH is facilitating this type of search by providing the most relevant browsing content and links with respect to a rich representation of user characteristics (such as preferences, history or prior knowledge). AH is assisting a user's information exploration by creating or adapting the composition, navigation and presentation across hypertext and hypermedia using e.g. link creation, link/content hiding or link/content annotation.

PIR is based on the standard Information Retrieval model, which is traditionally focused on the retrieval of documents that are relevant to a unitary query. While PIR extends this model by taking into account historical interactions, the paradigm remains that of finding the most relevant documents for a single user query. This fundamental underpinning of PIR makes such systems particularly suitable for the general information access paradigm of searching by query, where it is assumed that a user can express their information need in a relatively precise user query. However, this assumption is not always correct, especially in cases where users are uncertain about their actual information needs or the correct query terms to express these needs.

Despite AH and PIR stemming from such distinct information access paradigms, both fields effectively strive towards similar aims of providing the most personally relevant information to an individual user. This has led to many conceptual commonalities between the techniques of both fields, which can potentially complement each other in order to better assist a user's information exploration. The aim of assisting such searches is also actively researched in the area of Exploratory Search (Marchionini, 2006). This field distinguishes itself from standard search through the following 6 characteristics (White and Roth, 2009): 1) multiple query iterations, 2) open-ended information needs, 3) close coupling with task context, 4) combination of browsing and focused searching, 5) possible collaboration of multiple people and 6) advanced system evaluation with respect to learning, insight, task outcomes and system utility. One could thus argue that a hybrid of both AH and PIR systems could potentially be characterised by this search paradigm. By analysing and contrasting the various AH and PIR techniques, this chapter is able to provide a set of affordances that such a hybrid system's components require in order to improve user assistance across Informational and Exploratory search.

There are some other personalisation research fields that have also grown in prominence in the last 5 years, e.g. recommender systems and social search systems. For recommender systems, there are three common categories: content/feature-based recommender systems, collaborative/social recommender systems, as well as hybrid systems, which combine techniques from the former two (Burke, 2007). Content-based systems make use of the features associated with items as well as the interest rating a user has given to them. Such systems typically make use of statistical techniques commonly found in Personalised Information Retrieval systems, treating recommendations as a personalised classification problem (Burke, 2007). Collaborative filter based recommender systems are less content-oriented and users are typically presented with documents or items that are recommended by users with similar interests (e.g. Amazon⁶ recommendations). The neighbourhood of such peer users is typically calculated based on users' rating histories. Similarly, social search systems typically use such collaborative filtering techniques on user interest ratings in order to identify documents that were of interest to similar peer users. With the growth of online communities, these techniques might become increasingly powerful for future adaptive and personalised search.

However, since this thesis is focusing on information retrieval, presentation and content reasoning rather than the recommendation of items, such systems are not further explored in this chapter. Additionally, this thesis is not focusing on social and collaborative approaches to content personalisation, such as techniques presented in (Schafer et al. 2007).

For the purposes of analysing where adaptation can influence the search process, it is possible to characterise search adaptation technologies into three general stages: query adaptation, retrieval adaptation and result composition/presentation. First of all, the query adaptation step entails the analysis of an original query and its adaptation based on the available user information. Secondly, the retrieval algorithm can be adapted to incorporate personalisation features that tailor the retrieval to provide results that are deemed more relevant to the specific querying user. After the adaptive retrieval has produced personalised results, this information can be adapted further by composing and presenting the results in a coherent and personalised manner. Since the retrieval stage mostly provides unstructured results, these results need to be composed into a structured form that is most suitable for the querying user. This composition ideally

⁶ <http://www.amazon.com>

guides the user through the results and possibly provides a structure that helps the user explore the information in the most effective and efficient way. In addition to the result composition, the actual presentation can be adapted to suit the particular user.

The remainder of this chapter analyses each of the search adaptation stages, namely query adaptation (section 2.2), adaptive retrieval (section 2.3) and adaptive composition and presentation (section 2.4). The key approaches and techniques are identified and presented using examples from the research literature. Moreover, the different approaches are compared and critiqued broadly across their objectives, models, algorithms and scalability. The overall findings and the potential for a hybridisation approach are presented in section 2.5.

2.2. Query adaptation

The adaptation of user queries has been explored from many angles in several distinct research fields. The classic approach to the elicitation of an information query consists of a short interaction involving a user specification of a set of keywords. Although in theory these terms denote the user's original informational need, they are often short (typically consisting of only two to three keywords) (Spink and Jansen, 2004) and possibly ambiguous or incomplete due to common natural language problems such as homonymy and polysemy (Krovetz, 1997). This is reinforced by users generally providing low commitment to search interactions and having overly high expectations with respect to the search system (Jansen, et al., 2000). Since this is the main information entering an information retrieval engine, several attempts have been made to enhance the original query to more accurately reflect the user's perceived intention.

In a survey about semantic-based approaches, Mangold (2007) provides an overview of the different types of adaptation that can be applied to user queries. It is argued that most adaptation techniques are used to overcome the problem of ambiguity. The types of adaptations identified are: manual query modification (a user reformulating a query), graph-based modification (a user navigating a graph, which in turn generates new queries), query augmentation (expansion), query trimming and query substitution. It is argued that different techniques may perform better for different purposes (e.g. query expansion increases precision whereas query trimming increases recall), with some

systems now using a combination of different techniques. The survey below provides a summary of the implemented approaches and techniques.

Classical IR solutions for query adaptation have mainly consisted of statistical query expansion techniques, with the earliest systems performing Relevance Feedback (RF) from explicit user relevance judgments (Rocchio, 1971). In this approach, a user can specify which documents are relevant after an initial search query has produced a set of results. The most frequently occurring terms from the initially relevant documents are then used to expand the original query. This technique, albeit increasing precision and recall in research evaluations, is compromised by the fact that users are required to invest extra effort compared to regular search (since users have to go through multiple iterations to refine their query). Given that most users put very little effort into search, this could be considered as too cumbersome. A technique called Pseudo-Relevance Feedback (Xu, 1996) has automated this process by automatically selecting the top ranked documents as relevant information sources to perform relevance feedback. No user intervention is required, although it is arguable if this technique can still be regarded as personalised adaptation. Another popular statistical method for choosing expansion terms uses the concept of co-occurrence. For a certain term x in a user query, terms that frequently co-occur with x in the document base are shown to be excellent query expansion candidates (Kim and Choi, 1999). This technique is also evident in web search engines in the form of query suggestions provided either while eliciting the query or after an initial search has been conducted.

More recently, researchers have focused on improving the personalisation in relevance feedback by making use of users' search history and context. Koutrika and Ioannidis (2005) propose to mine search queries and relevance judgements into a user profile. This information can then be used to perform query disambiguation using query expansion based on the mined relationships. Logical operators such as AND, OR, NOT can be used to issue a new personalised query for the particular user. In Pitkow et al. (2002), a user bootstraps a user profile by importing bookmarks into the Outride tool. The bookmark links are then mapped to the top 1000 ODP⁷ categories in order to create a weighted user profile. Additionally, a recent account of user interests is kept by continually updating the profile with a user's click through history. When issuing a new query, the terms are compared against the user profile in order to select appropriate expansion terms. The mapping of websites to ODP categories is achieved by first

⁷ The Open Directory Project (ODP). <http://dmoz.org>

training a keyword-based classifier on a set of documents listed for each category (see Figure 2-1).

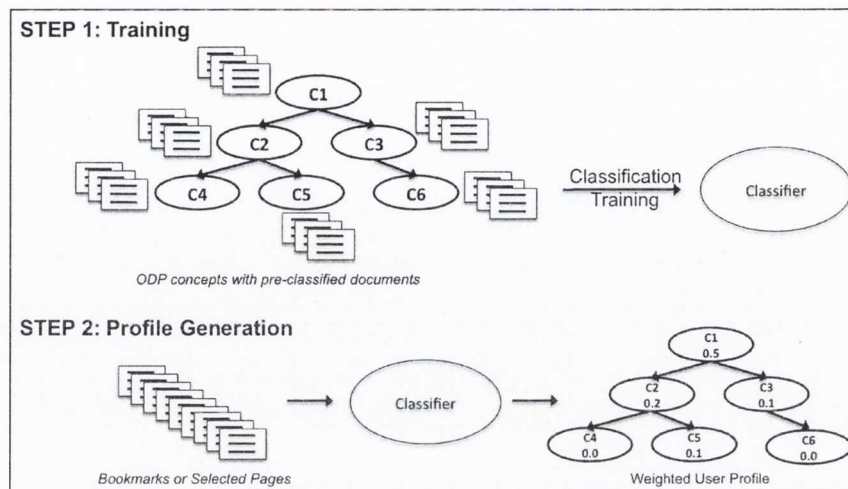


Figure 2-1. User profile generation using ODP categories

The user profile is then generated by classifying the bookmarked websites (i.e. the textual content) to ODP categories. This classification may also result in weighted similarity scores, allowing a more fine-grained representation of user interests (Micarelli, et al., 2006). Similarly, in Teevan, et al. (2005) the corpus for relevance feedback is not chosen from top ranked documents of an initial query, but from a user profile. This information is gathered from users' personal desktop applications (e-mail, word processing, etc.), as well as their browsing and search histories. It is argued that this information better denotes the short and long term user preferences. A similar approach was also taken by Chirita, et al. (2007), who were using a Personal Information Repository to choose expansion terms. Additionally, the number of query terms is adaptively chosen based on the query clarity. This query feature is calculated using both the scope of the query (i.e. its document coverage), as well as the query language model diversion from the document collection language model. Experiments show that further improvements can be achieved by adding such adaptive functionality.

Another form of implicitly derived relevance feedback mines users' interactions with search engines or other information systems to promote or reorder results returned to users. The outcomes are purely statistical and do not require any content-specific labelling or processing of items, and also do not require any explicit statement of interest from the user – the relevance feedback is purely a by-product of the user's other activities. An example of such implicit relevance feedback occurs in the use of clickthrough data and coselection data – a click (or clickthrough) is where a user selects

a specific result, which is interpreted as positive feedback on the relevance of the selected result to the search term, and can thus be used for result reordering. A coselection is where two or more results are selected from the same search, which implies not just relevance of the selected results to the search term, but additionally implies relevance of the selected results to each other. When aggregated, it is possible to use coselections to disambiguate the search term (Truran et al., 2005).

In Calegari and Pasi (2008), it is proposed to create a fuzzy ontology from users' search query histories and the documents stored on their personal workstations. Relationships are represented as simple numerical relationship strengths, which are displayed visually when a user issues a new query. The user may then choose related terms in order to expand the original query. Similar ideas are explored in (Ahn et al., 2010), where named entities are extracted from an initial set of retrieval results. These entities are then organised by their prominence and displayed along with the result list. The user can then select one or more named entities as expansions to the original query terms. Figure 2-2 gives an example of this interface, where the user has selected “Franz Schausberger” (left) and “Salzburg” (right) in addition to the original user query “train fire”. Such efforts move closer to ideas explored in Adaptive Hypermedia, which inherently make use of semantically rich conceptual models of user interests.

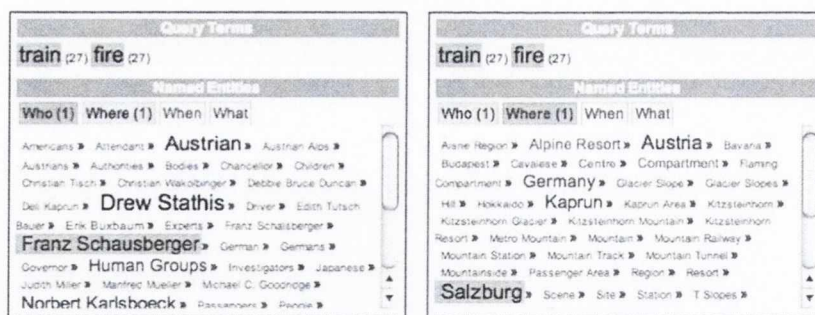


Figure 2-2. User interface for manual selection of additional query terms (Ahn et al., 2010)

Although the field of Adaptive Hypermedia typically does not explicitly involve user queries (as it has historically been based on browsing as opposed to searching), several related ideas can be found in the use of semantic retrieval techniques. Just as Adaptive Hypermedia systems use metadata, concepts and conceptual relationships to drive their adaptation, semantic-based retrieval makes use of such metadata and rich conceptual models to adapt the retrieval process (see section 2.3.2). In addition to this retrieval adaptation, several semantic-based systems also make use of such models to produce

initial query adaptation. A survey by Bhogal, Macfarlane and Smith (2007) reviews the use of ontologies for query expansion and argues that corpus-independent knowledge models can be used to provide disambiguation using the implicit semantic knowledge embedded in such models. Many surveyed systems make use of the Wordnet ontology synsets (Miller, 1995). This general ontology can be used to find related concepts, such as subclasses and superclasses, as well as synonyms to perform query disambiguation. Navigli and Velardi (2003) make use of ontologies in several ways by using a multitude of semantic relationships, e.g. hyperonymy (is-a), meronymy (has-a), similarity, etc. They compare the effects of several expansion techniques using general knowledge bases and conclude that a general improvement occurs over unexpanded queries. They argue that shorter queries benefit more from expansion due to the ambiguity associated with them. They also propose emergent semantic similarities between concepts by searching for “common nodes”. This technique chooses words for expansion based on the fact that they have similar synsets to the original query terms. The similarity is calculated by searching for nodes that both the original and the expansion node have in common.

Rocha, Schwabe and Aragao (2004) take an even more sophisticated approach by applying spreading activation to the underlying knowledge base. A query is expanded by a set of terms that have been “activated” in the knowledge graph due to their semantic relationships with the initial terms. This activation spreading is using ontology relationships, coupled with different weights attributed to each of these relationships. Depending on the particular domain or user preferences, stronger weights could be attributed to particular relationships in order to provide a “personalised” activation set. An important ontological relationship that is only rarely explored is antonymy (opposite of). Burton-Jones, Storey and Sugumaran (2003) argue that by including this in an expanded query with a preceding logical negation, query disambiguation can be enhanced further than by simple hyperonymy-based solutions. An even richer use of the semantic reasoning capabilities in ontologies is proposed by Linckels, Sack and Meinel (2007), where user queries are translated into Description Logic (DL). Although their work concentrates on searching over a closed knowledge base, it is worth noting that the semantic power of DL in ontologies could be used in several ways to reason about the expansion/reduction of concepts, as well as the logical operators separating the different terms.

It is interesting to note that the above methods all assume that an initial user query is adapted to form only one new query. However, Radlinski and Dumais (2006) propose to generate a set of related queries from an initial query by using a large sample of query logs. This set of queries is then used to generate results that can be reranked during the composition stage. The query choice is calculated by analysing the probabilities of which queries from the log are most likely to follow the initial query based on past usage patterns. Similarly, Glover et al. (2001) generate a set of modified queries from the initial query in their Inquirus2 system and then submit these to a selection of search engines (see Figure 2-3).

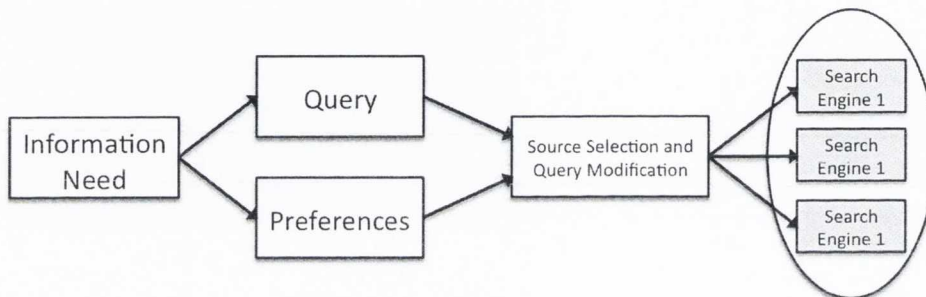


Figure 2-3. Personalised meta-search engine (Glover et al., 2001)

The modified queries however do not take into account semantics related to the query, as they rather attempt to broaden the search by appending terms such as “links” or “resources” for particular search types. The type of words added to the queries depends on a “need category” specified by a user, which denotes a certain query intent. Additionally, the type of search engine, as well as particular query constraints (such as recency) are chosen with respect to the particular need category. Similarly, in Kumaran and Allan (2008), the authors compare Interactive Query Reduction (IQR), Interactive Query Expansion (IQE) and a hybrid Selective Interactive Reduction and Expansion (SIRE). SIRE selects the top five sub-queries and the top five expansion queries and presents the results to the user, who can then guide the system implicitly to preferred queries. Results show that SIRE outperforms a baseline system as well as both IQR and IQE.

2.2.1. Summary and Critique

The analysis of query adaptation and personalisation shows a broad range of techniques that have been developed using either statistical or semantic approaches.

Classic Statistical techniques are shown to focus on bulk document analysis and keyword similarity in order to produce expanded queries. Since the focus lies on large document collections, no additional knowledge bases or user modeling components are required to apply these techniques. It can therefore be applied to large open corpus domains with high efficiency. This constitutes the biggest advantage of this type of system and it is therefore the most widely deployed approach on a web scale.

Although users can perform relevance feedback on initial sets of ranked documents, these systems do not keep a persistent record of particular user preferences, nor do they attempt to model the document space in a structured way. The personalisation aspect of these techniques is therefore limited, as user queries are regarded as ad hoc interactions. The relevance feedback never gets aggregated into some larger comprehension of relevance to the search term and thus each user needs to provide feedback for the same query. Moreover, these techniques only consider query expansion, not taking into account that query reduction or substitution might be more effective in certain cases.

Personalised Relevance Feedback techniques have attempted to acquire additional personal information of users in order to perform improved statistical similarity measures. Furthermore, not just query expansion but also query trimming and logical operator additions are considered for the adaptation of queries. Since the main focus still lies on statistical similarity measures, these systems can also be applied to large open corpus document bases. Personal information repositories as well as ODP categories are used as simplified knowledge bases that improve the personalisation effectiveness compared to classical statistical models. Again, the most important advantage of these systems is their current applicability on a web scale, since the knowledge base creation and document classification make use of efficient statistical similarity measures.

However, user models are still represented as simple keyword or keyword-relationship vectors, which contain little semantics in order to infer personalised query adaptation strategies. Moreover, these techniques do not perform semantic domain model reasoning, nor do they contain strategies for result diversification.

Semantic Techniques introduce the notion of a semantic domain model, which incorporates rich relationship information as well as reasoning rules. Query disambiguation is achieved using these relationships in conjunction with explicit user feedback.

However, the notion of a user model has been widely unexplored by these systems, leaving the adaptation process on a non-personalised level. Additionally, queries are not diversified, as the main objective of these systems is to simply disambiguate a single query. Furthermore, semantic-based techniques have largely been confined to closed corpus domains due to the current reliance on semantically rich models. Since it is difficult to create these models automatically, most systems have been built manually in order to query a small to medium sized digital library.

A solution to the non-adaptive nature of semantic retrieval techniques would be to use more advanced Adaptive Hypermedia techniques such as user modelling and personalised strategies in order to better adapt queries. However, the scalability issue described above would still persist, as both fields suffer equally from high model and metadata requirement costs.

Meta-Search systems have explored the diversification of search results by rewriting/generating sets of multiple queries, using either statistical similarity measures or particular expansion strategies. The term ‘meta search’ is used here to denote both systems that generate multiple queries to the same underlying search engine and systems that send queries to several different engines. Since all these systems do not just focus on providing improved single queries, they capture a greater breadth of search results. Additionally, by relying on statistical similarity measures, these systems are able to operate on large open corpora. Personalisation occurs after a user provides relevance feedback or an intent elicitation. Using this information, systems can adapt the resulting queries more precisely towards particular information needs.

However, the notion of expansion strategies introduced in Glover et al. (2001) has been confined to simple hardcoded rules that are applied for every user with the same query. Firstly, this leads to a scalability issue for the creation of new rules, since they each have to be created manually. Secondly, they have yet to be personalised in order to adapt the strategy to the particular querying user. Furthermore, diversification strategies generally do not make use of a knowledge base or semantic user model in order to reason about adaptive query expansion/trimming.

2.2.2. Comparison across Query Adaptation techniques

In conclusion, the analysis of query adaptation approaches reveals a variety of techniques that have been applied in order to improve initial user keyword queries. A comparison of these techniques across a number of properties can be found in Table 2-1.

	Classic Statistical	Personalised Relevance Feedback	Semantic Techniques	Meta-Search
Objectives	Query Disambiguation/Personalisation	Query Disambiguation/Personalisation	Query Disambiguation	Result Diversification
Knowledge Base	Document Set	Document Set, Personal Information Repository (PIR)	Semantic Knowledge Base	Document Set
Document Scale	Large Open Corpus	Large Open Corpus	Small to Medium Closed Corpus	Large Open Corpus
Model Requirements	none	ODP categories	General/Domain Ontology	Diversification Strategy
User Involvement	Relevance Judgements	Relevance Judgements	Explicit User Feedback	Relevance Feedback, User intent elicitation
User Model	none	PIR Relationships, Weighted ODP categories, Named Entities	none	none
Adaptation algorithm	Statistical Similarity (Keyword-based)	Statistical Similarity (Keyword-based)	Semantic Reasoning	Statistical/Strategy-driven
Query Modification	Weighted Query Expansion	Query & Logical Operator Expansion/Trimming	Query Expansion/Substitution	Query Substitution/Expansion/Trimming, Multiple Query Generation

Table 2-1. Summary of Query Adaptation Techniques

The *objective* of most techniques has been confined to simply disambiguating user queries, based on the assumption that the sparse set of initial keywords does not contain enough information to provide precise results. Furthermore, apart from several personalised relevance feedback systems, most approaches apply adaptation on a non-personalised level, largely ignoring user intent identification and processing.

Additionally, only a few systems move away from the generation of single expanded queries towards query diversification using multiple queries. A clear gap can be identified in the exploration of combining the different objectives, which would consist of an adaptive strategy for choosing appropriate and personalised adaptation techniques.

In terms of *knowledge base* and *model requirements*, there is a clear distinction between statistical and semantic techniques. Statistical solutions generally only require a large document set in order to calculate (personalised) relevance scores. Recent systems make use of personal information repositories and ODP categories in order to provide additional result classification and personalisation. This independence from complex domain models constitutes a great advantage over semantic solutions, which can only be applied to domains where such models have been created. Although the full combination of both scalability and domain model reasoning would constitute the ideal solution, a compromise could be reached by using statistical document set analysis together with semi-formal domain ontologies.

As mentioned before, the *document scale* has been the main focus of statistical IR systems. This constitutes a clear advantage over semantic retrieval systems, as it allows them to be applied to any large open corpus. The challenge of such systems hence lies in improving semantic query adaptation techniques in order to improve their current scalability issues. The incorporation of statistical similarity measures into semantic based retrieval systems as well as Adaptive Hypermedia systems proves necessary, as the loosening of metadata requirements could provide more scalable, lightweight solutions.

In order to perform personalisation, systems either require *user involvement* such as ad hoc relevance judgements or they create User Models in the form of more complex weighted sets of keyword relationships and category classifications. Users are often required to manually judge sets of relevant documents or they are asked to pick expansion terms based on the mined relationships. The more semantics that are encoded in the user profile, the more is it possible for systems to infer personalised query adaptation strategies. Personalised relevance feedback systems have been successful in mining large amounts of basic information from users' search histories and personal workstations. An improved solution would index this information using increased semantic relationships in order to represent users' preferences and requirements more precisely. Additionally, this would allow a system to apply personalised adaptation

strategies that could infer the right adaptation techniques for the particular user and context. By contrast, the sole dimension of adaptation that is currently being used is user interests, mined from previous interactions. Especially the user intent of a query would be of great use to a system in order to create a customised strategy for search result diversification. Currently, only the system provided by Glover et al. (2001) provides a basic implementation of this feature by requiring a user to explicitly choose a query intent.

Adaptation algorithms have been clearly divided into statistical similarity measures and semantic reasoning techniques. The clear advantage of statistical approaches lies again in the large scale applicability. However, it would be of great advantage if this strength is combined with domain and user model reasoning. Both methods do not exclude each other, they could rather be regarded as complementary. Especially in the case of insufficient semantic knowledge being available, statistical techniques could provide basic query adaptation functionalities.

Traditionally, the actual *query modification* focused solely on performing (weighted) query expansion. However, research has since been addressing a variety of adaptation techniques, each of which perform improvements over unmodified queries depending on the user and the context. An improved solution would adaptively choose the right modification based on the current user and domain model state. For example, a query could be both trimmed and expanded for a novice user in order to provide more guidance about the general domain. An initial adaptive algorithm has been presented by Chirita et al. (2007) where the number of expansion terms is chosen depending on the ambiguity of a query. This technique could be expanded by not only reasoning about the amount of keywords, but also about the type of modification that should be applied.

2.3. Retrieval adaptation

After an initial user query has been issued and possibly adapted, a retrieval engine is responsible for the retrieval of appropriate content. An adaptive personalised retrieval system targets the specific user information need and adapts to particular user preferences and context. Two main categories of techniques can be distinguished depending on the retrieval being based on statistical methods or on metadata-based algorithms.

2.3.1. Statistical methods

This category of algorithms is typically concerned with processing high volumes of data. Rather than replacing established IR/Web search algorithms, adaptive/personalised components are often attached as slight modifications or simply combined with the original search results. Ranked list scores are the preferred output of this type of algorithm and therefore the main focus of adaptive IR retrieval lies on improving the rank of documents that are relevant to the particular querying user.

An example of a modified retrieval algorithm is presented in Tanudjaja and Mui (2002). The ODP web taxonomy is used to capture user preferences, which then influence the HITS algorithm. HITS in its original form estimates the authority and hub values of a page solely using the given link structures (Kleinberg, 1999). However, the modified version gives more weight to pages that are related to positively rated pages (based on relevance feedback in the user's previous searches). Since each page can be mapped to the ODP taxonomy, relationships can be established by searching up and down the ODP tree for nodes that were previously explored. Similarly, Haveliwala (2003) provides a modified version of the popular link analysis algorithm PageRank. PageRank makes use of link information in order to provide a measure of popularity and authority of a page within a given set (Brin and Page, 1998). The presented modified version precomputes a set of topic-sensitive PageRank vectors for 16 ODP categories. At query time, the system first calculates the similarities of the query to the topics. Using these similarities, the system then adaptively calculates a linear combination of the topic-sensitive vectors for result ranking.

Furthermore, users' past searches can be analysed to disambiguate a query, since the terms might be matched to several categories. Additionally, the authors propose to use the context in which the query was issued. For example by highlighting the search term on a website, the user provides valuable context, namely the complete web page where the term was chosen from. Another example of modifying an established retrieval algorithm is shown in Teevan et al. (2005). The well-known ranking technique of BM25 is modified in order to incorporate user interests. BM25 is a probabilistic ranking function that includes document and query term weights and which incorporates relevance feedback information (Croft et al., 2009). The proposed modified version performs a new type of personalised Relevance Feedback, with the user information being gathered from rich personal Desktop information. Therefore, it is able to infer

relevance more accurately, since the re-ranking is based on documents that a user has actively interacted with.

In contrast to modifying established retrieval algorithms, many techniques calculate personalised result scores, which are then combined with original retrieval scores in order to determine the final result ranking. For example, the Wifs system in Micarelli and Sciarrone (2004) reranks initial search results (from Altavista) using similarity calculations between a user model and the returned documents. The user model is constructed using relevance feedback and contains terms that occur in the favourably rated documents. More specifically, terms that also occur in a manually constructed Terms Data Base (TDB) are considered as user model topics, whereas co-occurring words (found in the document but not in the TDB) are connected to the topics as co-keywords. Similarly, documents are represented using the occurring terms from the TDB, as well as non-TDB words co-occurring in the document. Several relevance calculations are then applied between the user model and the set of analysed documents in order to rerank the original results list.

Similar to the modified ranking algorithms, many combined ranking techniques make use of directory structures in order to rerank initial results. For example in Speretta and Gauch (2005), user profiles are constructed by mapping past queries and selected documents to ODP categories. The results for a newly issued query are analysed and mapped similarly to ODP categories and a similarity score is generated between the result documents and the user profile. This score is generated by multiplying the relative weights of concepts in the result documents and the user profile. The higher this similarity score, the more the results are deemed to be personally relevant to the user. By contrast, Daoud et al. (2010) use graph-based ranking models, whereby the similarity score is calculated as a combination of the minimum common supergraph and the maximum common subgraph of the result documents and the user profile (similarly based on ODP categories). By combining the similarity score with the original search rank, result documents that are more relevant for the particular user consequently appear higher up in the list. A similar approach is taken in Pitkow, et al. (2002), where an original result list is re-ranked based on a user profile. However, this profile has been created by categorising bookmarking links (imported into Internet Explorer by the user) onto 1000 ODP categories, as well as categorising over time the search results selected by the user. By contrast, Xiang et al. (2010) only use the immediate context of a query,

i.e. only successive queries (and their associated categories) are considered to be related (and consequently used for a personalised/contextualised score).

In addition to creating a weighted concept hierarchy, Liu, et al. (2004) associate different weights between particular keywords and the detected categories. For example, if a user interested in both “cooking” and “computers” has previously issued a query “apple” to retrieve “cooking” related documents, but not to retrieve “computer” related documents, the user profile will have a higher weight for “apple” in the “cooking” category. Therefore, this system has a higher degree of granularity over the system by Speretta and Gauch, since it takes into account more refined user preferences. An even more sophisticated classification is proposed in Stamou and Ntoulas (2009), where a topical ontology is created using ODP categories in conjunction with the Wordnet and SUMO ontologies. Users’ past queries are mapped to categories using several methods, including ontology traversal. The discovered topics are then used as a user profile during the rank combination phase in order to create a personalised document ranking.

2.3.2. Metadata-based approaches

The second type of adaptive retrieval systems heavily relies on rich models and document annotations. Semantic-based retrieval systems as well as Adaptive Hypermedia systems are built on the power of semantic knowledge engineering, aiming to achieve the vision of the Semantic Web (Berners-Lee et al., 2001). In terms of retrieval, techniques and algorithms utilised in Adaptive Hypermedia have been inherently conceived to provide adaptive result retrieval (in addition to the adaptive navigation and presentation functionalities presented in section 2.4.2). Since many prototype systems have been developed for the field of E-learning, most of the research has focussed on the personalised retrieval of learning content. However, the proposed techniques and algorithms are not exclusively applicable to this particular domain.

The notion of domain and user models has been proposed by the earliest Adaptive Hypermedia prototype systems, such as Interbook (Brusilovsky, et al., 2004) and AHA! (De Bra, et al., 2003). The domain model typically represents a conceptual view of the underlying domain, containing information about concept hierarchies, attributes and relationships. This model is independent from the underlying content and represents a more high-level model of the subject domain. For example, in an e-learning domain this

model may contain various high-level topics to be covered within a subject, including relationships such as prerequisite requirements between topics. Interbook proposes the idea of mapping each document to a set of outcomes and prerequisites. It is argued that storing adaptation-specific information in external models assists the adaptive retrieval by allowing reasoning engines to infer which documents are relevant for the user. For example, the set of outcomes and prerequisites of a document allow a reasoning system to provide a user with a set of documents that should be visited before and after the current document. Additionally, using an overlay user model, Interbook is able to infer a student's knowledge state for each of the domain concepts, which allows a personalised delivery/omission of information. Furthermore, the notion of a learning goal is proposed, which defines a particular sequence of documents in order to guide a student through the material. Similarly, the content adaptation in AHA! bases the inclusion/exclusion of fragments to be shown to the user on the state of the user model in relation to the domain model. Depending on previous user visits to certain pages, AHA! checks the suitability of a particular fragment in order to provide a student with personalised pages. In an improved version, called GALE (Smits and De Bra, 2011), it is shown that domain concepts can be distributed over several servers in order to decentralise the adaptive functionality. In Aroyo et al. (2004), it is proposed to incorporate external fragments into the AHA! system by making use of Information Retrieval (IR) results. The authors make the assumption that a resolver ontology describing the search space can be mapped directly to the domain ontology. By doing so, the standard fragment inclusion/exclusion techniques can be performed on the domain concept level as in the case for the regular AHA! system. However, no experimental prototypes have been developed to evaluate these possibilities.

In the KBS hyperbook system (Henze and Nejd, 2001), documents are again linked to external models. However, as opposed to Interbook, the KBS hyperbook documents are only indexed with domain concepts, with the inter-concept relationships being defined in a Knowledge Model. Again, a user model is used to capture the actual knowledge of a student in order to compare this to the knowledge required to understand the topic in question. The system adaptively retrieves the set of concepts that the user should learn about first (the prerequisites). During the adaptive retrieval, all Knowledge Items that should be learned by the user are marked based on users' knowledge of the items and their prerequisites. The actual documents are then selected based on the document-concept indices that have been created a priori. All adaptation therefore occurs during

the concept retrieval stage, rather than the content retrieval stage. It has to be noted that any document can be incorporated into the KBS hyperbook system as long as it is indexed to the concept space. The authors argue that a drawback of the system is that the learning dependencies are explicitly encoded into the concept space, although these relationships might differ for different scenarios. Depending on the intention of the knowledge engineer who designs the domain model, different assumptions might get encoded into the knowledge model. For example, different instructors might teach a certain subject using a different teaching strategy, leading to different prerequisite requirements encoded into the knowledge model. The authors acknowledge that this goes against the idea that a knowledge model should be independent of the particular teaching strategy.

The idea of selecting learning objects according to a particular strategy is shown in Farrell, Liburd and Thomas (2004). Initial XML search results are mapped to concept domain topics, for which further learning objects are retrieved in order to form a coherent learning path. The adaptive retrieval is based on the original query and the statistics collected for each topic during a mapping stage. Moreover, users can indicate their desired course duration depending on the time that is available to them. A drawback of the system is the fact that the particular retrieval strategy is encoded into the system rather than separated from the adaptive engine. In Conlan, et al. (2002) and Conlan and Wade (2004), a multi-modal, metadata-driven approach is proposed in order to provide this separation of concerns. Most importantly, the approach introduces an additional model called the narrative model, which encodes a set of generic strategies for presenting concepts. For example, this model can encapsulate an expert's knowledge of a domain and therefore provide guidance through appropriate course material. The implemented APeLS system executes the narrative by consolidating models in an Adaptive Engine. In terms of retrieval adaptation, the narrative contains the rules for which concepts should be selected, how they should be sequenced and which candidate content group should be considered for content retrieval. All adaptation hence occurs on the concept-level, enabling the narrative to be independent from the actual content. The adaptation rules can be based on the information that is available in any of the models, for example user prior knowledge, media type preferences, history, cognitive style, etc. (each held in the Learner Metadata Repository). Narratives can be implemented using several technologies, such as rule-

based (e.g. Drools⁸) or script-based languages (e.g. Javascript). Since all adaptation occurs on the concept level, the closed corpus could be expanded by taking a similar approach to the KBS hyperbook system. Furthermore, the idea of a narrative as a way to guide a user could be applied to different retrieval adaptation techniques in order to ensure coherent strategies.

The use of ontologies as concept domains is proposed in semantic-based systems such as the Personal Reader Framework (Henze, 2005). In this system, recommendations for learning resources are again based on the current learning progress of the user. Standardised metadata annotations are used in order to infer which learning object should be recommended/retrieved. Also, an alternative recommendation service is proposed, which could be based on the keywords that describe the objectives of the learning object in an ontology. This would allow different course materials to be used in the system. It is therefore argued that different recommendation services might be suitable in different situations. Furthermore, the use of more generalised ontologies for user observations and for adaptation makes it possible to share models among different applications. In Dolog, et al. (2003), it is proposed to fully move adaptive systems towards Semantic Web technologies in order to enable an adaptive Semantic Web. Due to the standardised formats, interoperability between applications would hence be greatly facilitated. It is argued that adaptive retrieval can be based on common logic-based languages, such as First-Order Logic. Since Semantic Web technologies are inherently focussed on reasoning capabilities, rule-based languages such as TRIPLE (Sintek and Decker, 2002) can be employed to reason across distributed metadata. The enhanced expressiveness of RDF⁹ and OWL¹⁰ allow the creation of comprehensive domain, user and adaptation models in order to perform semantic retrieval adaptation. In Linckels, Sack and Meinel (2007), learning objects are adaptively retrieved exclusively using such inferences of Description Logics, which is possible due to the semantically marked up documents. In Tran, et al. (2008) it is even proposed to have one domain ontology that encompasses all the different aspects of Adaptive Hypermedia, such as the user or the task. Additionally, a model for the specification of adaptation rules upon this ontology is suggested to capture the adaptive behaviour in a declarative manner.

⁸ <http://www.jboss.org/drools>

⁹ <http://www.w3.org/TR/rdf-schema>

¹⁰ <http://www.w3.org/2004/OWL>

However, the use of such detailed concept, metadata and rule modelling requirements has confined many of the presented techniques to rather small-scale applications. More “lightweight” semantic techniques are proposed in Fernandez et al. (2008), where documents are semantically annotated using a combination of keyword frequencies, as well as semantically related documents’ keyword frequencies. Furthermore, their ranking algorithm makes use of a combination of conceptual ranking together with standard keyword ranking in order to lift the semantic techniques to a potentially large scale. While such semantic retrieval approaches do not have any explicit representation of user needs, there exists great potential to utilise or modify such techniques for adaptation and personalisation. An example of such personalised semantic techniques is shown by Cantador et al. (2008), where a semantic user model is created from the concept annotations that are associated with documents viewed by the user. Additionally, the user model concept weights are constantly updated depending on the frequency of user interactions with associated documents. The document ranking then makes use of these different concept weights in order to provide a conceptual ranking that compares the user preference vector to the document metadata vector. By combining a standard keyword search score with such a conceptual ranking, search results can be shown to be more personally relevant to querying users. However, the single axis of adaptation is the adaptation towards previously shown interests, resulting in an improved ranked list only.

2.3.3. Summary and Critique

From a statistical IR point of view, it has been suggested to either modify traditional ranking algorithms or to combine original scores with an additional personalised score. Metadata-based approaches can be divided into techniques that stem from the fields of Adaptive Hypermedia and the Semantic Web.

In the case of *modified statistical ranking* algorithms, traditional scoring formulas are either extended with additional parameters or they are biased towards more personally-relevant information. Personal relevance judgements, as well as directory structures are often employed to categorise both documents and users in order to calculate similarities across them. Additionally, personal desktop information is utilised to gather larger volumes of data that can be used to find personally relevant search results. Statistical keyword-based measures are used to find similarities between different sets of

documents and user preferences. This constitutes a clear advantage of these types of systems, since the indexing of documents and their classification can be applied on a large scale.

However, the surveyed classifications occur on a very broad scale, as they often employ only a small subset of ODP categories. Additionally, the user model is confined to a simple set of (weighted) keywords, which is lacking additional semantics in order to infer more refined adaptation strategies. Especially in the case of an initial short keyword query, this type of technique does not take into account that additional documents beyond the provided set of terms are relevant to the user. There is a clear lack of a strategy or narrative for adapting the results to suit the particular information intent or need of the user.

Similar conclusions can be inferred for techniques that combine an original ranking function with additional personalisation features (*combined statistical ranking*). These approaches often rerank an initial set of results according to previous user interactions or relevance judgements. ODP categories and users' query histories are used in order to create weighted user profiles. This fully automated process is again scalable for large open corpus collections, which constitutes the main advantage of these techniques.

However, due to the inherent reliance on an original, non-personalised set of results, it can be argued that a large set of personally relevant results is neglected in the initial retrieval. Even if the result list is reranked successfully, these systems are not diversifying the results according to users' actual information needs.

In the field of *Adaptive Hypermedia*, rich domain and user models are used in order to retrieve personally relevant information. Due to the fact that many systems have been developed for the field of e-learning, user models have often focussed on modelling a user's knowledge of the domain. Together with specific user preferences (e.g. learning styles), this knowledge model is used by systems to adaptively retrieve documents that a user should examine in order to fill a knowledge gap. The adaptation usually occurs on a conceptual level, with relationships between concepts being defined in the domain concept model. These systems allow a multi-dimensional adaptation that can utilise multiple user attributes for personalisation.

However, since these systems operate on a concept level, a mapping has to be created between documents and concepts. This has restricted most AH applications to a closed

corpus domain, with certain open-corpus solutions requiring considerable indexing effort. Furthermore, the retrieval of concepts is usually not initialised by a user query, as most educational systems focus on delivering a personalised course rather than a query response. This leads to the open issue of how to apply the developed adaptation techniques to the often sparse sets of user keyword queries.

Semantic techniques fully move towards semantic technologies such as ontologies for domain and user models. Additionally, adaptation rules can be encoded into ontologies using description logics and new query languages can be used to reason across semantic knowledge bases. Interoperability is the clear advantage of these techniques, as such systems might be able to share domain, user and adaptation models. Such techniques can be used for both adapting a conventional retrieval algorithm (i.e. perform PIR), as well as for reasoning about results and adaptively retrieving additional resources (i.e. perform AH).

Again due to the high construction costs for such systems (especially the concept-content indexing), closed corpus domains have dominated most of the research in this area. However, more “lightweight” semantic techniques that require less metadata and that include some statistical elements seem to improve this scalability issue, with the downside of focusing on user interests only.

2.3.4. Comparison across Retrieval Adaptation techniques

Due to the inherent differences between statistical and metadata-based retrieval approaches, various techniques have been developed to add adaptivity and personalisation to the retrieval process. A comparison of the surveyed techniques can be found in Table 2-2.

In terms of *document indexing*, systems that use modified ranking algorithms or that combine original search results with a reranking module utilise standard keyword frequency measures. This can be calculated automatically and can therefore be applied to large document sets. For an Adaptive Hypermedia system, documents have to be mapped to domain concepts, which is often done manually. IR systems therefore clearly outperform Adaptive Hypermedia applications in terms of *scalability*. However, current indexing techniques do not cover additional document characteristics, such as its suitability for different types of users and contexts (e.g. novice/expert, time constraints,

etc.). These types of document indices can often only be created manually or semi-automatically. Adaptive Hypermedia has its strength in handling these different dimensions of adaptivity that could be applied across a document base.

	Modified Statistical Ranking	Combined Statistical Ranking	Adaptive Hypermedia	Semantic Techniques
Document Index	Keyword Index	Keyword Index	Concept-Content mapping	Concept-Content mapping
Scalability	Large Open Corpus	Large Open Corpus	Closed Corpus/ Small Open Corpus	Closed Corpus/ Small-Medium Open Corpus
Domain Model	ODP categories	ODP categories	Bespoke Domain Concept Model	Domain Concept Ontology
User Model	ODP categories, Personal Information Repository	ODP categories, Keyword History	Overlay Knowledge Model, User Preferences	Overlay Knowledge Model, User preferences
Metadata Requirements	Automatic topic categorisation	Automatic topic categorisation	Rich content metadata, Concept-content mapping	Rich content metadata, Concept-content mapping, Ontology-encoded knowledge bases
Adaptation algorithm	Modified HITS/ PageRank/ BM25	Reranking of initial statistical result, Category similarity calculation	Rule-based application of strategy/narrative	Semantic reasoning, Combined Statistical-Semantic Ranking

Table 2-2. Summary of Retrieval Adaptation Techniques

The *domain model* of statistical systems most often consist of ODP categories in order to classify and personalise ranking results. These techniques have the clear advantage of being applicable to large collections of documents without human interaction. However, the strength of Adaptive Hypermedia techniques lies in the more detailed specification of domain models using either bespoke concept models or general ontologies. Although these models are more labour intensive to create, they allow adaptation strategies to reason over which concepts should be included in the retrieval. Especially in combination with a semantic user model, these domain ontologies can infer concept dependencies that are not captured adequately by keyword similarity measures.

Similar conclusions can be drawn from the surveyed systems in terms of *user models*. Statistical techniques rely on the machine processing of either ODP category classification or personal information repository mining. Weighted keywords denote user preferences, with minor keyword relationships being established using co-occurrence. The strength again lies in the processing of large sets of documents, which can lead to the mining of potentially large amounts of personally-relevant keywords. However, a semantically richer user model allows systems to infer which additional relevant information should be presented to the user in order to address the current information need. Additionally, the surveyed AH systems explore a multitude of dimensions of adaptation, using an increased number of user variables (e.g. prior knowledge, learning styles, time available). By contrast, the sole dimension explored by statistical techniques is based upon user interests that have been mined from previous search interactions.

As a consequence, the *metadata requirements* for statistical IR techniques can be computed efficiently, as techniques such as information extraction and topic categorisation can be fully automated. The disadvantage of Adaptive Hypermedia (including semantic-based techniques) lies in the reliance of not only marked up documents, but since the reasoning occurs on a concept level, a concept-to-content mapping needs to be created. Since current systems require a rigid annotation and indexing of these resources, manual effort is usually required to produce the desired quality of metadata. A better solution would loosen these requirements in order to provide more lightweight reasoning solutions. Additionally, this would allow the integration of (semi-) automatic IR indexing techniques in order to process larger open corpora.

The *adaptation algorithms* used by the statistical systems either integrate personalised features directly into the ranking function or they combine an initial score with a separate personalisation score. In both cases, they are applicable on large document bases, relying mostly on either term frequency measures or category similarity calculations. As mentioned before, combined score techniques only rerank an initial set of results that might be missing valuable information. It is therefore advisable to include personalisation features directly in the ranking formula in order to find personally relevant documents across the whole collection. Adaptive Hypermedia techniques make use of rule-based strategies or narratives that capture particular adaptation techniques. Using these rules, it is possible to extend the retrieval beyond an

initial set of user defined keywords by examining which additional resources address a user's information need. An improved solution could apply these techniques to an open corpus space by combining the strategy model (or semantic reasoning in the case of semantic techniques) with loose concept indexing of large document corpora.

2.4. Adaptive Composition & Presentation

The set of results returned after the adaptive retrieval stage can be personalised further during the composition and presentation stage. Due to the variety of algorithms and methods presented in the previous section, the type of output results differ significantly from system to system. Therefore, an array of techniques has been developed to perform different types of adaptive result composition and presentation. Again, these methods can be divided into statistical and metadata-based techniques, depending on the algorithm that generated the results.

2.4.1. Statistical techniques

Traditionally, ranked lists have been the preferred method of displaying statistical information retrieval results. This is reinforced by the fact that most research in the IR field, including result adaptation and personalisation, is concerned with improving single-valued relevance scores, which can only be displayed in the ranked list format. Although this score might be calculated using several features/algorithms (including personalisation features), a single aggregated score is generally calculated in order to simply compare the different values for ranking. However, recent research has attempted to provide alternative composition and presentation techniques in order to provide more personalised, adaptive and diversified results. A selection of such systems is surveyed below.

First of all, the composition (merging) of search results from a set of multiple search queries has been explored by several researchers. The idea behind such techniques lies in the broadening of search results in order to either focus a user's search towards more precise information needs or to improve the performance of personalised reranking techniques. For example, in the metasearch engine Inquirus2 overviewed earlier, Glover et al. (2001) use the results retrieved by various search engines and compose these into a single ranked list. As opposed to typical metasearch engines, they not only consider

the titles, summaries and URLs for the rank merging, but the whole pages returned by each engine. The ranking is based on multi-attribute utility theory, which takes into account several factors, depending on which need category was chosen by the user. The different preferences are used in an additive value function, which combines the different metadata fields that are available. For example, the indicated preference for “current events” would put a 60% emphasis on “TopicalRelevance” and 40% on “DaysOld”. Similarly, Radlinski and Dumais (2006) also propose the diversification of search results through the merging of multiple result sets. However, they propose to generate the set of multiple queries by determining related queries from a large sample of query logs. In combination with reranking techniques proposed in Teevan, et al. (2005), the diversified results are shown to provide improved personalised rankings.

The idea of retrieving diversified search results is also proposed in the meta-search systems in Sushmita, Joho and Lalmas (2010) and Thomas et al. (2010). However, these systems make use of an “aggregated search interface” in order to compose and visually present a more diverse set of results. Similar to the diversification techniques above, an initial query is sent to several information sources in order to retrieve diversified search results. The resulting documents are then not simply combined into a single ranked list, but they are displayed in a separate panel for each information source on the same “aggregated” results page (similar to Yahoo alpha¹¹). Experimental results show that these prototype systems enable users to look at more diverse results, select more items to complete their tasks, and that they are generally perceived to be superior to a standard ranked list system. The idea of such systems hence lies in the immediate visualisation of more diversified search results, as opposed to just attempting the improvement of a single merged result ranking.

Another approach of using statistical methods to adaptively compose and display results is the concept of clustering. In Xu, Jin and Lau (2009), a user query is sent to a third-party web search engine to retrieve N number of results. This set is then clustered into different topics using standard document clustering techniques. Following this step, the main topics are allocated a display panel in the visual interface. The size and location of each panel depends on the size and importance of the search results contained in each cluster. A user can then either click directly on one of the search results from a cluster, or expand a particular cluster to display the full results of the chosen topic. Yippy¹²

¹¹ <http://au.alpha.yahoo.com/>

¹² <http://www.yippy.com>

takes a similar approach by categorising search results into folders and subfolders. A user can then expand a certain folder/subfolder to refine the search, therefore making the interface adaptive to user interactions. In Truran, Goulding and Ashman (2005), multiple clicks (“co-selections”) on a set of search results are interpreted as indicating mutual relevance. By mining such relationships, their system is able to aggregate search results of ambiguous queries into a set of clusters that can help users sort through the different query senses. Although this usage of collective intelligence has so far focussed on non-personalised aggregations, it is worth considering the creation of class/cohort specific systems using this technique.

Some research has been conducted in improving the traditional result presentations by providing the user with increased information about the retrieved resources. In Psarras and Jose (2006), an improved summarisation system is proposed, which performs adaptive query-biased summarisation. These summaries are presented with the traditional ranked list and are shown to improve users’ relevance judgement during result browsing. The system is implemented as a recommendation portal, which adaptively presents relevant documents to a user based on previous searches and result visits. In White, Jose and Ruthven (2003), the “WebDocSum” interface similarly provides users with an improved summary in the form of a summary window. When a user moves the mouse over one of the query results, this window displays a summary for the document. In addition to the usual fields such as title and summary sentences, it provides the user with the number of outlinks on the page, the first non-text object and the document size. Similarly, in Joho and Jose (2008) an evaluation of 4 different search result presentations is performed. A baseline ranked list is compared to (i) a system that presents top ranking sentences along with each result, (ii) a system that shows a thumbnail image of each result document (i.e. a screenshot of the actual document) and (iii) a system that presents both top ranking sentences and screenshots with each result. Although neither of the two research works above describe adaptive result presentation, it is argued that differences could be noticed among users with different search experience. More specifically, it is noted that less experienced users might benefit from the systems that provide extra information for the search results in order to make better relevance judgements. Furthermore, the search interface should be made adaptive to the particular task, context and user experience to offer the right and appropriate assistance at any given time. In related research, Villa et al. (2009) provide an adaptive “aspectual” search presentation that allows users to model search subtasks.

For each aspect, users can have a separate panel with its own history, undo history, current query, search results, etc. This adaptive presentation allows for the completion of complex information needs, which require users to search for multiple aspects within the overall task.

While the techniques and approaches above attempt to better organise, compose and present an initial set of search results, Bhavnani et al. (2003) propose to use domain knowledge from experts in order to develop actual search strategies that can help novice searchers find information more effectively and efficiently. They developed the idea of Strategy Hubs, which provide initial selection categories, which in their case are related to medical conditions. A user can initially choose from a selection of diseases, followed by a selection of subcategories, such as “Treatment” or “Diagnosis”. In a second step, the hub provides specific search strategies about how the user should find information related to certain topics. Additionally, for each of the strategy steps, links are provided for reliable sources that are known to provide good information. The strategies consist of a series of sequenced steps that have been identified by experts. In terms of presentation, a dual-frame design has been chosen, which displays the different steps in an upper frame and the actual content pages in a lower frame. It is argued that this design provides a consistent user interface, supporting novice users in their perception of the overall strategy. Since novices have greater difficulties in identifying sub-goals when searching for comprehensive answers, the strategy hub can support users by guiding them towards a more structured way of searching.

2.4.2. Metadata-based techniques

In the presence of rich metadata, faceted rankings have become a popular way of composing and presenting IR results in a more easily comprehensible manner. Such systems allow users to search for information through the specification of more refined attributes than just simple keywords (Yee, et al., 2003). For example, in Figure 2-4 a user has refined the image search according to “Location: Asia” and “Shapes, Colours, and Materials: fabrics”, resulting in a narrowed down set of image results. Such metadata attributes are typically added manually to individual items in a collection, although they can be extracted automatically to a certain extent. Such a faceted way of ranking search results is also proposed in Teevan, et al. (2008), where facets are described to “represent a dimension that can be used to organise information”. It is

argued that an adaptive process could choose between different facets depending on the user task and context, as well as the document domain. By selecting multiple rank facets (using different document attributes), a user could adapt the result presentation towards a more personalised view of relevancy scores. This idea is also proposed in Tvarozek and Bielikova (2007), where only selected metadata fields are used to show/rank the most relevant attributes of the search results. Furthermore, the authors propose an adaptive version of their faceted browser, which provides automatic facet selection based on user preferences, global attribute relevance and inter-attribute relationships. Additionally, facets can be adaptively ordered (based on their estimated relevance), annotated (e.g. using tooltips), and recommended based on particular restrictions (e.g. IT companies being recommended to an IT consultant). Similarly, Zhang and Zhang (2010) propose to recommend document facet-value pairs to users and to incorporate the selected values into the retrieval models. Experimental results show that for a corpus of semi-structured text documents, a non-boolean retrieval model performs more effectively.

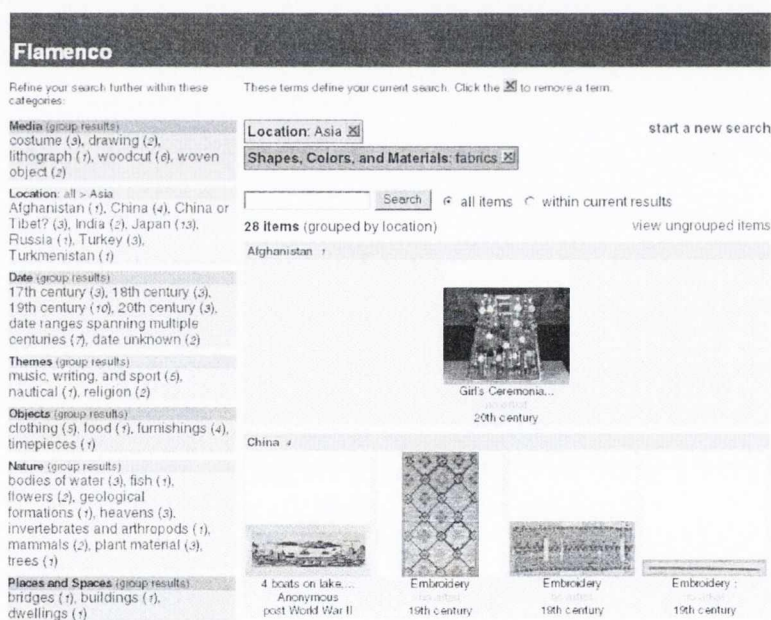


Figure 2-4. Example of a faceted search interface (Yee, et al., 2003)

In contrast to traditional IR systems, the field of Adaptive Hypermedia (AH) has been inherently focused on providing users with adaptive result compositions and presentations. Due to the availability of rich user and domain models, coupled with bespoke content metadata, ideas in this field focus around composing appropriate information sequences and navigations and displaying these using personalised presentations.

As mentioned earlier, these systems have generally focused on delivering educational material to students in order to provide adaptive e-learning courses. For example, as mentioned in the adaptive retrieval section, the KBS Hyperbook system described in Henze and Nejd1 (2001) performs concept-level adaptation by making use of domain and user models. The selected concepts are sequenced according to prerequisite requirements, creating an order of links that guides a student towards the next best document to view. It is not left to the student to sort through a ranked list of documents, as the system advises an appropriate path through the document space through link sorting/hiding. Since the actual documents have been indexed with concepts from the domain model, any open corpus document can be sequenced appropriately. Similarly, the idea of providing appropriate sequences of documents is shown in Farrell, Liburd and Thomas (2004), where XML search results are mapped to topics and then sequenced according to particular concept domain rules (e.g. to teach more basic information first). Additionally, the actual objects within topics are sequenced according to an “Instructional Role Sequence” (for example introductions are sequenced before concept procedures and conclusions). However, the sequencing service is embedded into the adaptive system rather than implemented as a separate adaptation model. Also, no personalisation is provided apart from the possibility to perform query refinement.

The generation of personalised learning sequences is taken further in the APeLS system (Conlan et al., 2002) (Conlan and Wade, 2004), where the separation between the core adaptive engine and the sequencing service is proposed in the form of a narrative model. This narrative reflects a didactical ordering that can be specifically adapted to the current user task, context and preferences. By applying this adaptive narrative, the system can provide a personalised result sequence based on the particular user knowledge and preference levels, hence guiding the user through the document space (again through link generation, link sorting, etc.). Although the system has been initially conceived to work over a closed corpus, it would be possible to index and consequently integrate open corpus documents, as in the case for the KBS Hyperbook.

The ideas of composing and sequencing information have also been explored using a combination of ontologies and description logics. For example, Karam, et al. (2007) assemble learning objects by inferring the best “composition flow” using the current user knowledge state and the domain ontology. Description logic is used to solve the “concept-covering problem”, which corresponds to the knowledge need of the current

user. Similarly, in Geurts, et al. (2003) ontologies are used in order to provide a "structured progression" through retrieved results. The concept of narrative units is proposed, which are used to construct the complete result presentation in a structured order. Each retrieved semantic unit has rules associated with it, which dictate the information that should follow the current unit. As a consequence, after the application of the complete set of rules, the result is a structured progression through a semantic graph.

In addition to composing such adaptive navigations, AH systems often use adaptive presentation techniques such as link adaptation in order to provide users with presentational hints. For example, by colouring or hiding selected links, Smits and De Bra (2011) provide hints to users about the relative suitability of particular knowledge items. With growing knowledge of a user, different links either become available or get coloured to symbolise their "readiness" to the user. The knowledge modelling in such applications relies on rich domain models, which can be used as a basis for overlay user models in order to infer the suitability of particular content. In Hsiao, et al., (2009) and Hsiao, et al., (2010), a user's progress through a course-test system is tracked in order to provide adaptive link annotations. Such annotations provide students with hints about which task to try next and also visualise how often a certain quiz has been attempted already. In Figure 2-5, the target icon in the menu (left) presents the growth of student knowledge (shown by the number arrows) and the relevance of the topic to the current course goal (shown by the colour intensity of the target, ranging from faded to strong intensity). Such techniques are shown to lead users to attempt more course tests and to have higher success rates, as adaptive link annotations seem to have a high motivational effect in such applications.

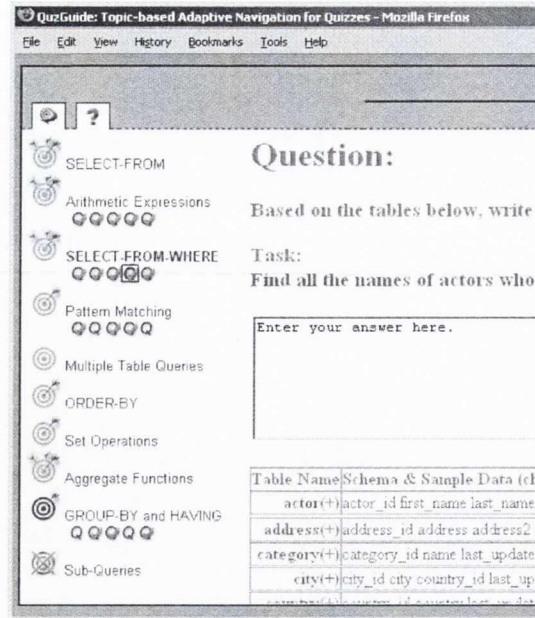


Figure 2-5. Example of link annotation (Hsiao et. al., 2009)

In Jovanović, et al. (2006), sets of knowledge items are grouped into an annotated tree of links and link annotations are provided to show users which documents are most appropriate based on prerequisite requirements. Upon selecting one of the links, the system then generates a new assembly of learning content based on the selected topic.

By contrast, rather than showing a user a full sequence of documents, the systems in Henze (2005) as well as in Smith-Atakan and Blandford (2003) consist of an adaptive presentation in the form of related information that are apart from the currently viewed content. The former presents the current best pages on the left of the current window, whereas the latter system, called ML Tutor, provides these in a separate window. In Brusilovsky, et al. (2004), The Knowledge Sea II system provides a user with a map where similar documents are placed in adjacent cells. It provides social navigation by using visual cues based on an individual user's browsing history combined with all other system users. The popularity of a particular document is highlighted both on an individual and on an overall level, guiding a user towards popular documents that he/she has yet to visit. Additionally, users can provide annotations, such as positive or negative feedback, which can further help fellow users find interesting documents. However, as noted by the authors, although the system is very efficient in adding open corpus documents, it is lacking a strong navigation support, especially for an educational system.

Further adaptive presentation techniques include scaling, where important information is highlighted through increasing the size of relevant content, or stretch text, whereby less relevant content is only represented by placeholders (Tsandilas and schraefel, 2004).

2.4.3. Summary and Critique

The adaptive composition and presentation of information has been studied by both fields of Personalised Information Retrieval and Adaptive Hypermedia. Each approach is trying to overcome the information overload problem by grouping, sequencing and presenting documents in a coherent manner.

In the field of Personalised Information Retrieval, most research has generally focussed on adapting the result compositions and presentations by reranking initial search results. The presented techniques move away only slightly from the current ranked list paradigm by grouping results into clusters or by ranking the results according to particular facets. Additionally, the current features of ranked lists (e.g. result snippets) have been adapted in order to provide more personalised result presentations. As in the previous sections, statistical document analysis techniques provide this improved visualisation. Document clustering and snippet/summary improvements can be applied on a large scale, making these techniques attractive to be applied in current web search.

The concept of result diversification and aggregated search moves away slightly from the single query and single ranked list paradigms in order to present users with a greater breadth of search results. However, neither technique makes use of adaptive strategies in order to choose the right type of query diversification and information source for a particular user, task or context.

Additionally, the notion of information sequencing has yet to be addressed in order to assist the information searcher more adequately. Search strategies or procedures are not supported, leaving users having to filter through large ranked lists in order to satisfy their information need. Also, the presented systems do not take into account that users might have previously acquired particular information about a certain subject, which would ideally decrease the future relevancy of documents that cover this part of the knowledge space.

On the other hand, due to their inherently adaptive behaviour, Adaptive Hypermedia systems address this user guidance using various approaches. Concepts and content are composed into sequences of coherent (learning) paths, which assist the user in finding additional documents for their initial information need. By making use of user models that contain current preferences and knowledge levels, this path can be adapted to form a personalised response for a particular user. Such systems also have a strong focus on applying adaptive presentation techniques, effectively guiding users through content using adaptive hints.

However, due to the high reliance on metadata and concept-content indexing, most systems have been confined to small closed corpus spaces. The only system, which does not require careful metadata indexing is the Knowledge Sea II system by Brusilovsky et al (2004). However, it is argued that the social navigation component in this system does not provide the same quality in terms of user guidance (Brusilovsky and Henze 2007). It is therefore important to carefully balance the issues of scalability and navigation support in order to provide an open corpus system with adequate levels of user guidance and personalisation.

2.4.4. Comparison across Adaptive Composition & Presentation techniques

The survey of the different systems reveals a variety of techniques to add adaptive behaviour to the composition and presentation stage. A comparison of the different techniques and approaches is provided in Table 2-3.

In the field of Information Retrieval, the *adaptive behaviour* has focused on slightly evolving the current ranked list paradigm by adding either cluster visualisation, faceted ranking or improved snippet/summary generation. Although these techniques are easily applicable on a large scale, they do not provide any of the advanced adaptation techniques that are present in Adaptive Hypermedia systems. Current AH approaches address the very important aspect of concept/content sequencing, which provides more guidance for querying users. Furthermore, the presentation adaptation of AH systems often provides additional visual cues (such as link colouring or hiding) to the user in order to show the personal suitability of particular documents. This adaptive behaviour represents one of the true strengths of AH systems and could therefore provide

excellent extension possibilities to current IR systems in order to make use of their bulk processing in a more personalised manner.

	Adaptive Composition		Adaptive Presentation	
	Information Retrieval	Adaptive Hypermedia	Information Retrieval	Adaptive Hypermedia
Adaptive Behaviour	Faceted Ranking, Search Aggregation, Result Clustering	Adaptive navigation, Concept/content sequencing, Link ordering, Map-based indexing	Adaptive result summarisation, Adaptive cluster presentation	Link annotation, Link colouring, Map-based presentation, Stretch text, Scaling
User Involvement	User intent elicitation, Facet preference selection	none	Cluster selection Aspect creation	none
User Model	none	User Knowledge/ Preferences	Result summarisation preferences	User Knowledge/ Preferences
Metadata Requirements	Intention-facet relationship model	Concept-content index, Keyword similarity	none	Concept-content index, Keyword similarity
Adaptation algorithm	Statistical document/keyword analysis	Rule-based (ECA), Social-based, Semantic reasoning	Statistical document/keyword analysis	Rule-based (ECA), Self-organised social mapping
Scale	Large open corpus	Closed Corpus/Small-Medium open corpus	Large open corpus	Closed Corpus/Small-Medium open corpus

Table 2-3. Summary of Adaptive Composition and Presentation Techniques

In terms of personalisation, IR and AH systems have both increasingly investigated the concept of facet preferences in order to provide more suitable rankings for a particular user.

User involvement is still the most valuable input for IR systems in order to infer precise and context-sensitive relevance scores, whereas AH techniques make use of *user models* to store a user's prior knowledge and additional preferences in order to predict appropriate information relevance. Although this addition would prove very useful in IR systems as well, it constitutes a great challenge to mine such information using standard data mining techniques. In order to retrieve this prior knowledge information from open web data, one would first need to map users' page visits to a domain model and consequently estimate a user's information gain about the appropriate concept. This

problem represents a fundamental issue in implicit feedback techniques, as a user's browsing activity does not necessarily correspond to learning about the documents' topics. Different approaches can be applied to address this problem. For example, one possibility is to assign suitably low value-adds for page visits in order to avoid overestimating a user's experience. Alternatively, rather than looking at a user's search or browsing sessions (page counts, word counts), another approach could employ public user profile mining. Once a user has been identified, user profile data from social networking sites such as Facebook or Twitter could be mined and consequently utilised within the personalised application (Abel et al., 2011).

The *metadata requirements* for the surveyed IR systems are comparably low, as the main focus has been on automatic document classification and result summarisation. This independence from markup data makes these systems very suitable for large-scale corpora such as the web. However, the emerging concept of faceted ranking will require considerable amounts of metadata to order documents according to particular preferences. From an AH perspective, this dependency on sufficient metadata has been apparent since the earliest systems. Documents need to be linked to domain concepts in order to reason about the suitability for a particular user. However, systems such as the Knowledge Sea system by Brusilovsky et al. (2004) have made initial steps towards a more open AH system. The compromise between reliance on metadata, user guidance and scalability has to be chosen carefully in order to provide the right level of adaptivity for the particular task and context.

As mentioned in the previous sections, the *adaptation algorithms* used by IR systems have built on statistical keyword similarity and clustering measures in order to provide result compositions. User relevance feedback is captured and added to the ranking formula in order to update the rankings/clusters. AH systems on the other hand have mainly made use of rule-based algorithms to compose suitable result flows. Ontologies are being used increasingly, which leads to the encoding of adaptation rules in formal languages such as description logics.

Since these rules are currently not being created automatically, a *scalability* issue arises again for AH systems. A solution to this problem would be to use high-level adaptation strategies that could be complemented by automatic IR processing capabilities.

2.5. Conclusions

The fields of Adaptive Hypermedia (AH) and Personalised Information Retrieval (PIR) have each recognised the challenge of adaptive and personalised information delivery. However, due to their inherent conceptual differences (presented in section 2.1), the techniques and technologies have varied substantially between the two approaches (described in sections 2.2, 2.3, 2.4).

Section 2.5.1 first presents several conclusions that can be drawn in terms of the user dimensions (i.e. characteristics) addressed by PIR and AH systems. Section 2.5.2 draws a number of conclusions about the different adaptation techniques that PIR and AH have applied to adapt to such dimensions. Lastly, in order to overcome some of the identified weaknesses, section 2.5.3 investigates the potential benefits of combining PIR and AH techniques in a hybridised approach.

2.5.1. User dimensions

In the field of PIR, adaptation and personalisation techniques have predominantly focused on the statistical analysis of historical usage and corpus patterns, using for example past queries, query refinements or user clicks (see sections 2.2.2 and 2.3.4 for a comparison of techniques). The analysis of such usage patterns has typically mined sets of user interests, which can then be used for narrowing future retrievals towards related information.

As discussed in sections 2.2.2 and 2.3.4, the major advantage of such techniques has been shown to lie in their scalability, as the algorithms are mainly focussed on automatic processing of large volumes of data. Search logs, personal information repositories, as well as directory structures such as the ODP have been used increasingly in order to personalise search results by categorising both users and documents. Some systems have even mined minor semantic relationships between documents, concepts and users by analysing query histories and consequent click behaviours, including e.g. co-selections on a set of search results such as in Truran et al. (2005).

However, it is unavoidable that such techniques may introduce substantial noise, especially due to the variety of user information contexts and (sometimes extemporary)

needs. One way to overcome such noise (e.g. misrepresentation of user interests) is through scrutability and user control, whereby users are able to manually review and adjust the current user model. Moreover, it is crucial to recognise a user's current task in order to provide not only personalised information, but also contextualised personalisation. Initial attempts have been made in order to tackle this issue, for example, Cantador et al. (2008) model short-term "context models" in a similar fashion to the more long-term user models (i.e. by mapping keyword vectors to semantic concepts), but using a strong decay factor to fade out older concepts. This context model is then compared to the user model in order to select the relevant subset of a user's long-term interests for personalisation. This ensures that a user's long-term interests are taken into account, while not over personalising results based on contextually irrelevant information.

The fact that PIR techniques typically base personalised relevance estimation solely on previous user interests also constitutes a very narrow focus of adaptivity articulation for such systems. Since users constantly interact with systems in order to fill particular knowledge gaps, it is crucial to consider that current information needs depend on a number of characteristics, such as a user's task or current knowledge state.

AH techniques have inherently focussed on capturing and using such additional dimensions of adaptation (e.g. prior knowledge, cognitive/learning style) in order to provide user assistance and guidance (see sections 2.3.4 and 2.4.4). Such multi-model techniques have mostly focussed on instrumenting the usage during a domain-specific user session and are therefore less ambivalent or susceptible to noise. Moreover, as AH techniques are inherently based around structured conceptual models rather than unstructured keyword histories, the aforementioned scrutability and user control can be achieved more easily through the development of model manipulation interfaces (Bakalov et al, 2010).

However, the rich conceptual modelling techniques of AH systems have typically been confined to narrow domains such as e-learning and it remains an open challenge to broaden some of the techniques to larger content bases. In particular, it may be necessary to also use statistical usage methods (similar to the presented PIR techniques) in order to track such information across open-corpus domains. Initial attempts have been made in the context of studying open-corpus novelty detection, whereby a user's domain knowledge is calculated using a knowledge accumulation method based on

previously viewed documents (Lin and Brusilovsky, 2011). The suitability of a new document can therefore not only be characterised by its similarity to user interests, but also by the estimated suitability to a user's current knowledge state. However, these novel techniques for opening up traditional AH knowledge modelling still need to be evaluated thoroughly in terms of their accuracy and usefulness in applied scenarios.

2.5.2. Adaptation techniques

In the field of PIR, adaptation techniques have most prominently focused on query expansion and result reranking. Such techniques have typically been based on statistical similarity measures using keyword-based user and document models (see sections 2.2.2 and 2.3.4). An array of different techniques has been proposed to expand, modify or trim initial user queries, as well as to bias retrieval algorithms towards statistical user models. The techniques have been shown to successfully retrieve focused result lists according to prior user interests and they have generally maintained the scalability of standard search systems.

However, the notion of adaptively choosing different query/retrieval adaptation techniques has been less explored to date. One area where such ideas have been actively researched is in the related field of Question Answering, where systems may involve query expansion based on synonyms, external thesauri or parsing a user's question with grammars of varying sophistication. The goal of such adaptive selections of multiple query adaptation techniques typically lies in broadening the potential pool of answers before proceeding to the actual answer extraction step (Hirschman and Gaizauskas, 2001). However, while such techniques could also be employed to broaden and diversify initial user queries, PIR systems typically revert to the presentation of their results using the conventional ranked list paradigm (see section 2.4.1).

This constitutes one of the main drawbacks of PIR approaches, as users are typically left having to filter through simple ranked lists (or possibly clusters of ranked lists) of potentially relevant/non-relevant documents. No explicit user guidance according to a strategy or narrative is provided across documents, a fact that is reinforced by the common IR batch-evaluation techniques that do not involve real-world users. This lack of narrative is one of the key distinctions between PIR and AH, and aligns with the distinction noted previously between searching (PIR) and browsing (AH) (see section

2.1). While PIR is almost entirely search behaviour, AH is characterised by browsing behaviour guided through a narrative defined previously by an expert.

By contrast, AH techniques have inherently focussed on providing such user guidance through external models to enable particular information seeking strategies. Adaptive result composition and sequencing are used to provide a flow of currently suitable information in order to provide “the right information at the right time” (see section 2.3.2). Moreover, as systems have typically used implicit queries (e.g. the statement of an intent/learning goal), AH techniques have focused on scoping the adaptive experience rather than adapting an initial keyword query. Many systems have been conceived for the application domain of e-learning, where user guidance can be provided by a domain expert (often a teacher/lecturer) through the encoding of a domain model. Adaptation strategies can then be applied on this model by defining rules upon relationships, such as prerequisite requirements. These rules can be applied in order to provide additional dimensions of adaptivity, such as localisation or context/task-based personalisation.

Moreover, AH systems also make use of a multitude of adaptive presentation techniques such as link colouring or annotation (see section 2.4.2), thereby providing a much more guided browsing experience. These presentation techniques present one of the true strengths of AH systems, as they can provide personalised hints without hiding information from users. This also highlights again the focus of AH systems on user navigation and interaction compared to the batch computation of ranked result lists.

However, due to the inherent reliance on refined concept indexing, most research has still been confined to very narrow domains, such as educational systems or cultural heritage libraries. Due to the closed nature of such systems, most research in recent years has focussed on how to move towards open-corpus domains. Lightweight solutions to concept mining, indexing, reasoning and adaptation are required, which make use of both the bulk-processing capabilities of (P)IR and the adaptation and personalisation approaches of AH. The Semantic Web field has also introduced such increasingly scalable solutions, although they have mostly had the downside of using their semantic capabilities for the purpose of only improving ranked lists using the sole dimension of user interests. Additionally, the Linked Data initiative promises a large-scale availability of structured data that could be used by AH and semantic systems. To date, over 32 billion triplets have been published already, which could help AH to

overcome some of its scalability issues. However, it remains to be seen how effectively these datasets can be used by adaptive systems.

2.5.3. Overall Findings and Complementary Affordances

Although both PIR and AH have attempted to address the same challenge of delivering personally relevant information, each approach has presented different strengths and weaknesses in terms of user dimensions and adaptation techniques.

First of all, PIR systems have been shown to typically only provide adaptation according to the narrow dimension of user interests. Moreover, the adaptation techniques have typically been confined to the simple alteration of search result rankings, which can lead to the previously described problems of over personalisation and low information diversity. In order to increase the breadth of adaptation capabilities of such systems, it is therefore perhaps crucial for PIR to embrace the notion of multi-dimensional adaptation that current AH systems provide. The AH notion of adaptive guidance in terms of result composition and presentation could also be beneficial to PIR systems in order to overcome the low user commitment in current search systems. By utilising adaptive composition and presentation techniques, PIR systems could potentially engage users into personalised search sessions and motivate them to subscribe to the notion of search as an interactive process.

Several weaknesses have also been shown for AH systems, most notably in terms of their strong reliance on rich metadata models for retrieving information. This characteristic of AH techniques has typically confined such systems to specific application domains such as e-learning. In order to overcome these weaknesses, AH systems need to embrace the power of statistical document analysis techniques that have been shown to successfully drive the adaptation in current PIR systems. Techniques such as keyword query expansion and selective information source selection could enable AH systems to provide their multidimensional adaptation across larger open-corpus domains.

However, it is only possible to provide these combined functionalities if the complete retrieval process is enhanced. Most research so far has focussed on providing adaptivity only during either the query adaptation stage, the retrieval stage or the composition/presentation stage. Very little attention has been devoted to providing a

unified adaptation approach, which could encompass all aspects of the task of information retrieval and delivery. However, it is crucial to align the different stages into a coherent workflow in order to enable personalised guidance during the information composition stage.

The development of “hybridised” systems could potentially combine different techniques and technologies in order to provide such a unified adaptation. The kind of affordances that a hybrid system may provide would thereby reach across i) query adaptation, ii) retrieval and iii) result composition and presentation.

First of all, multiple models can contribute towards the adaptation of the user’s query, including both AH-type metadata models about the user’s preferences and context, as well as PIR-type models of search histories. Moreover, AH strategies can be applied in order to choose between different types of query adaptation techniques depending on the different model states. This can also include the generation of a set of multiple queries of varying detail, complexity and source selection in order to maximise the diversity of results. In particular, this diversification can aid the later composition and presentation states by providing a broader range of information related to the current topic of interest. In addition to query adaptation, the retrieval of information can be adapted using Adaptive Hypermedia and PIR techniques. This again allows the final composition and presentation stage to better guide users across the various results. Finally, AH components can generate the navigation across the retrieved content based on the current model states.

However, there remain many challenges towards achieving such integrated adaptation and personalisation. In particular, it is of paramount importance that the various adaptation stages are coordinated in terms of end-to-end effectiveness. If such a harmonised combination of techniques is not taken into consideration, the various adaptation effects could potentially neutralise each other or in some cases even be detrimental towards the overall system performance. For example, it might be desirable that an application does not perform adaptation on the same characteristics twice, as this might skew or over blow the results too much. This could be the case if a system personalises a query based on a particular characteristic and then also performs personalised retrieval based on the exact same attributes. Similarly, if a system has a broad-type strategy across the result navigation (e.g. to give a user as much choice as

possible), it would not be advisable to perform too much focused personalisation during the earlier process stages.

The key step towards the successful application of the proposed approach hence lies in the joined-up thinking between the various adaptation characteristics and the understanding and managing of the trade-offs between techniques. Rather than arbitrarily combining multiple adaptation capabilities, it is crucial to develop an overarching strategy that takes into account the system's application context and goals. The various adaptation techniques then need to be coordinated according to this overall strategy in order to maximise the complementary affordances.

3 Initial Adaptive Open-Corpus Composition System

3.1. Introduction

As outlined in chapters 1 and 2, Adaptive Hypermedia (AH) systems typically focus on providing adaptive information compositions and presentations for formal or informal learners. Such compositions often adapt to multiple user characteristics and generally provide adaptive user guidance across the underlying content space. However, a common problem with such systems has been shown to lie in the need for handcrafted learning objects, as the material is typically sourced from a proprietary set of closed-corpus content. Moreover, the analysis of AH techniques has revealed a lack of adaptive response generations that satisfy informal user queries, since Adaptive Hypermedia systems have traditionally provided complete educational course compositions.

This chapter describes an initial investigation of the benefits and drawbacks in using such an AH system for generating adaptive information compositions that satisfy informal user queries. In particular, this chapter describes an initial adaptive composition system that (i) *provides adaptive compositions across open-corpus information* and (ii) *satisfies informal user queries*.

The remainder of this chapter is structured as follows. Section 3.2 first presents the author's contributions to the work presented in this chapter. Section 3.3 then describes the overall architecture of the underlying Adaptive Engine, which applies multi-model Adaptive Hypermedia design principles for the generation of the adaptive compositions.

An application of the architecture is shown in an e-learning prototype (section 3.4) and evaluated in an authentic learning environment in terms of educational benefit, user efficiency, satisfaction and motivation (section 3.5). By comparing the system to ranked-list based Information Retrieval prototypes, it is shown that the compositional approach motivates users to explore more resources while issuing the same number of queries. However, it is also shown that the composition system requires significant effort in order to integrate new open-corpus resources and that the query elicitation possibilities are very limited compared to standard search systems. Finally, section 3.6 concludes this chapter with a set of requirements for more advanced information composition architectures in order to alleviate the limitations that were identified. In particular, it is argued that additional open-corpus retrieval and adaptation capabilities are required in order to apply the adaptive compositional approach across large, dynamic and heterogeneous content bases.

3.2. Contribution of the author

In order to investigate the benefits and drawbacks of an AH-based open-corpus composition system, a number of pre-existing technologies were extended and combined to generate adaptive query responses. The author's contribution to the work presented in this chapter lies in the extension and usage of these technologies in an integrated process to create and evaluate a novel AH-based open-corpus query system. More specifically, the individual contributions of the author are as follows:

- The extension of an existing Adaptive Engine to handle semantic models (i.e. models that are specified using RDF¹³/OWL¹⁴) (see sections 3.3.2 and 3.3.3 for architecture descriptions).
- The specification and development of domain, content, user and adaptation models to be run in the Adaptive Engine (see section 3.3.1 for model descriptions).
- The integration of an open-corpus content harvester and annotation client to gather open-web content and its associated metadata (used as the content model) (see section 3.4.1 for a description of this process)

¹³ <http://www.w3.org/RDF/>

¹⁴ <http://www.w3.org/TR/owl-features/>

- The development of a web-based application, which generates personalised open-corpus responses to user queries through a number of adaptation process steps (using adaptation rules that are executed in the underlying Adaptive Engine) (see section 3.4.2 for a description of this process).
- The evaluation of potential usability benefits and drawbacks of the compositional approach, as well as an assessment of the technological limitations of such a metadata-driven AH-based architecture (see section 3.5).
- The development of two IR baseline systems for this evaluation (see section 3.5.2).

3.3. Architecture

The architecture of this initial composition system builds on the multi-model, metadata-driven AH architecture that was developed for the Adaptive Personalized eLearning Service (APeLS) (Conlan 2002). In this design, multiple models are consolidated in an Adaptive Engine (AE) in order to produce personalised information presentations.

Section 3.3.1 gives an overview of the models used in the application of this AH architecture in the initial open-corpus composition system (consisting of *Domain Model*, *User Model*, *Content Model* and *Narrative Model*). Section 3.3.2 describes the components and capabilities of the architecture, consisting mainly of *Strategy Interpretation*, *Model Control* and *Model Manipulation*. These described components also comprise semantic capabilities that extend the original architecture in order to allow the inclusion of semantic domain models. Finally, section 3.3.3 describes the technological architecture of the implementation.

3.3.1. Models

The *Domain model* represents a conceptual view of the underlying domain, containing information about concept hierarchies, attributes and relationships. This model is independent from the underlying content and represents a more high-level model of the subject domain. For example, in an e-learning domain this model may contain various high-level topics to be covered within a subject, including relationships such as prerequisite requirements between topics. The particular domain ontology developed for the experimental prototype is presented in section 3.4.1.

As a user can be characterised by multiple dimensions, the *User Model* is aimed at representing various user preferences, interests and context. In the typical Adaptive Hypermedia scenario, this model often consists of an overlay of the domain model, representing the relative knowledge of a user with respect to the domain model concepts. Furthermore, additional context and preferences may be contained in this model, such as the user's query intent or device capabilities. The particular user characteristics used in the experimental prototype presented in this chapter consisted of a user's query intent specification (during query elicitation, see section 3.4.2), as well as their prior knowledge in the subject domain (captured through a questionnaire, see section 3.5.3).

The *Content Model* contains the various metadata values held by the content, including for example the concept related to the content or its difficulty level. The degree of metadata depends on each individual data source in terms of amount and granularity. This model is typically added manually during the creation of the content itself and thereby restricts the integration of new open-corpus content. However, this restriction can be alleviated to a certain extent, since the metadata can be generated independently from the content creation stage. An example of such a separation will be shown in the implementation section (see section 3.4.1).

The *Adaptation Model/Narrative* describes the strategy by which the various concepts and content can be explored. It defines the overall "storyline" by adapting the composition and navigation in order to support particular objectives (as done in many textbooks where the authors give sample paths through the book for different levels of interest and/or background knowledge). For example, in an e-learning scenario the narrative could define a particular learning path according to a user's prior knowledge or learning preferences. This narrative could also define the inclusion/exclusion of particular types of content depending on a particular teaching strategy, focussing for example on highly example-based or more theory-driven teaching. The particular narrative process developed for the experimental prototype is presented in section 3.4.2.

3.3.2. Architecture Components & Capabilities

This section describes the overall Adaptive Engine (AE) architecture with its various components and capabilities. Figure 3-1 illustrates the separation of the main

components, namely *Strategy Interpretation*, *Model Control*, *Model Manipulation* and *Model Repository*.

As mentioned in the previous section, the AE architecture builds on the notion of an adaptation strategy for the encapsulation and definition of the adaptation rules/narrative across a number of models. The AE architecture allows for script-based, as well as rule-based adaptation narratives through different interpretation languages (see section 3.3.3 for technology implementation specifications). By separating the narrative model from the actual AE *Strategy Interpretation* capabilities, it is possible to flexibly change to a different narrative depending on different model properties or contexts. For example, depending on a particular user's information need or cognitive preferences, specific narrative models can be adaptively loaded and executed to best serve this user in terms of content selection and sequencing.

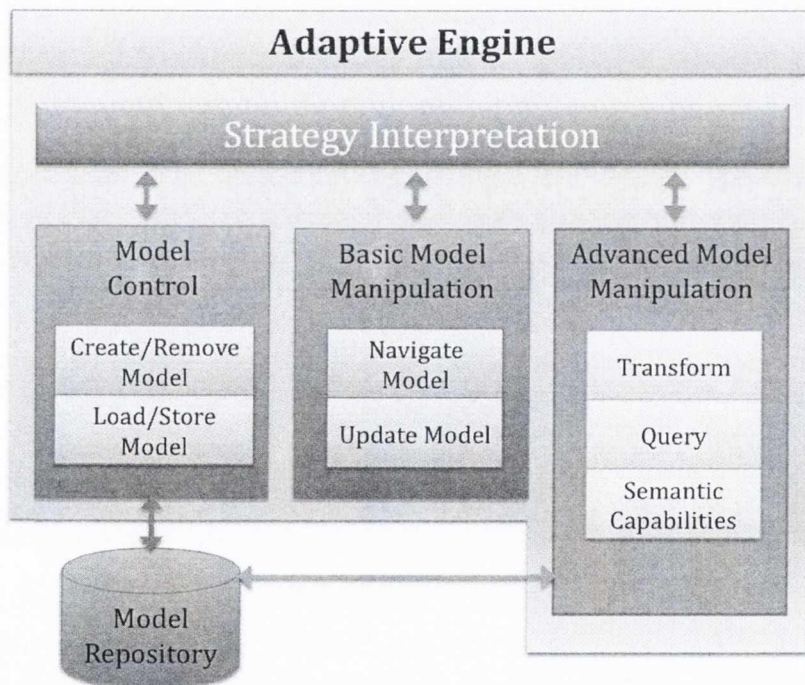


Figure 3-1. Adaptive Engine Component Architecture

The *Model Control* and *Basic Manipulation* capabilities enable the creation and manipulation of multiple models, allowing the creation of new model compositions (e.g. a model of a composed e-learning course), as well as model manipulations, such as navigating and updating model nodes and contents.

The *Advanced Model Manipulation* capabilities include the transformation of models from one form into another (e.g. a raw result model to a model that can be rendered by a web browser), as well as the querying of models using structured queries. As

mentioned in section 3.2, the original multi-model metadata-driven AE architecture has been extended to include *Semantic Capabilities*, allowing the integration, manipulation and querying of semantic data structures. More specifically, new libraries have been added to the engine to handle OWL/RDF-based models. This enables narratives to make use of the expressivity provided by standardised semantic technologies, including sophisticated domain/user modelling and triple-based querying across class hierarchies, attributes and relationships.

The number of models that can be used with this component architecture is unrestricted. The control and manipulation of the models is facilitated through an external model repository, which can consist of different underlying data storage technologies (see section 3.3.3 for technology details).

3.3.3. Technological Architecture

The technological architecture builds on the Adaptive Engine (AE) framework developed for the Adaptive Personalized eLearning Service (APeLS) (Conlan 2002). This framework consists of a set of Java libraries and implements the various components of the multimodel architecture described above. Figure 3-2 illustrates a technological view of the various component implementations of this architecture.

The *Strategy/Narrative Interpretation* capabilities are implemented for a variety of script- and rule-based languages. This allows the narrative developer to choose the most suitable type of language for expressing the desired adaptivity. Rule-based languages allow for the creation of rules that can adaptively fire in reaction to certain events or model states. Script-based languages allow developers to define a linear narrative to execute a particular strategy. The languages that a developer can choose from are Javascript (through the Rhino¹⁵ interpretation engine), Drools¹⁶, Jess¹⁷ and Jatha¹⁸ (Lisp implementation for Java). Each of these interpretation engines has the same complete access to the underlying AE capabilities, allowing narratives to make full use of the model control and manipulation functionalities. The narratives defined in these languages are stored as models and can be loaded and manipulated in the same manner as all other models held by the AE.

¹⁵ <http://www.mozilla.org/rhino>

¹⁶ <http://www.jboss.org/drools>

¹⁷ <http://www.jessrules.com>

¹⁸ <http://jatha.sf.net>

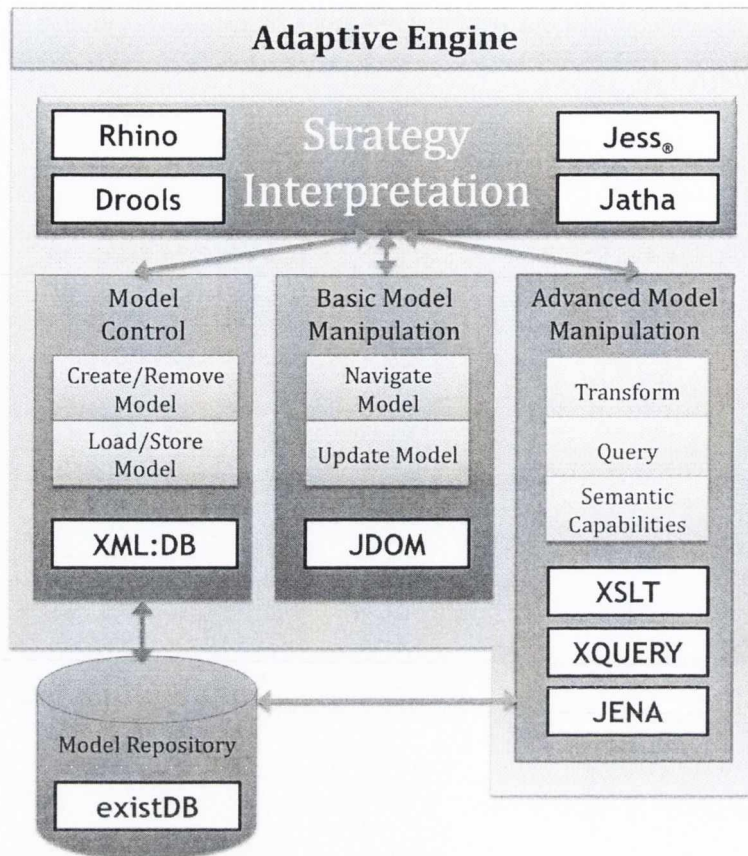


Figure 3-2. Adaptive Engine Technological Architecture

The *Model Control* and *Basic Model Manipulation* capabilities are implemented using the XML:DB¹⁹ API and JDOM API²⁰. The *Model Repository* in the current implementation can either consist of the local file system or a dedicated XML database (e.g. existDB²¹).

The *Advanced Model Manipulation* component includes transformation capabilities through an XSLT²² engine, as well as XML querying through XPATH²³ and XQUERY²⁴. As mentioned previously, *Semantic Capabilities* have been added to the original AE architecture in order to make full usage of semantic modelling and querying technologies. For this purpose, the Jena API²⁵ has now been integrated into the AE, allowing narratives to query ontological models through SPARQL²⁶ queries

¹⁹ <http://xmldb-org.sourceforge.net>

²⁰ <http://www.jdom.org>

²¹ <http://exist-db.org>

²² <http://www.w3.org/TR/xslt>

²³ <http://www.w3.org/TR/xpath>

²⁴ <http://www.w3.org/TR/xquery>

²⁵ <http://jena.sourceforge.net>

²⁶ <http://www.w3.org/TR/rdf-sparql-query>

(using the Jena ARQ²⁷ query engine). Moreover, custom functions can be added to the AE in order to use more advanced semantic capabilities such as ontological reasoning.

User Input/Output

In order to receive user queries and to display result compositions, AE instances are created and held by JavaServer Pages (JSP). These pages have full access to the AE and can therefore dynamically load and run narratives depending on user variables. After the final result model transformation (e.g. to XHTML), the output can be displayed to the user's web browser by simply including the composition into the JSP code.

3.4. Prototype Implementation

This section describes the application of the presented architecture in an open-corpus e-learning prototype. Section 3.4.1 first describes the prerequisite processes of the implemented open-corpus e-learning prototype, including the harvesting and annotation of the open-corpus content. Secondly, section 3.4.2 describes the complete adaptation process of generating an information composition for an informal user query.

3.4.1. Prototype Prerequisites

In order to apply the presented architecture in an open-corpus e-learning prototype, several processes need to be run a priori in order to i) harvest domain-specific content from the open web and ii) generate reasonably accurate metadata descriptions of the content. Figure 3-3 illustrates this content harvesting and metadata generation process.

²⁷ <http://jena.sourceforge.net/ARQ>

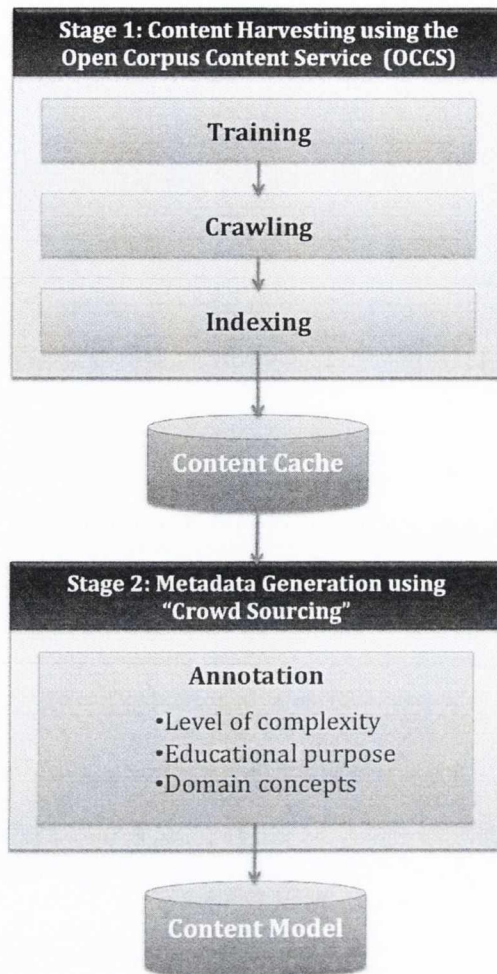


Figure 3-3. Content Harvesting and Metadata Generation process

Stage 1: Content Harvesting

In order to harvest domain-relevant content from the open web, an open-corpus content harvesting tool called *Open Corpus Content Service (OCCS)* (Lawless, et al., 2008) was used. This service takes a tool-chain architecture approach in order to discover, classify and harvest content from the World Wide Web. A focused web crawler is employed to conduct traversals of the web, seeking content in defined subject domains. The crawler functions by incrementally selecting a URI from among those scheduled and fetching the content located at the URI. The content is then classified to assess its relevancy to the scope of the crawl. Content classification involves the filtering of content for both language and subject domain. A text classifier is trained in advance of each crawl to generate a statistical model of the subject area. The OCCS then uses this model to ascertain the relevancy of crawled content to the scope of the crawl. Upon crawl completion, each harvested item of content is parsed and an index is created of the entire content cache.

Stage 2: Metadata Generation

In addition to a content cache, the metadata-driven AH architecture requires metadata about the collected documents in order to apply the desired adaptation and personalisation.

Although several approaches for automatic metadata generation exist (Reeve, 2005), most applications only capture a fraction of the types of document annotations that are required by Adaptive Hypermedia systems. For example, because the implemented prototype is focused on the domain of e-learning, several specific metadata fields are required, such as the level of complexity and educational purpose of documents. Consequently, a “crowd sourcing” approach has been taken in this initial study in order to retrieve this fine-grained level of metadata. An annotation tool has been used, which displays documents from the OCCS cache along with a predefined list of possible values for level of complexity, educational purpose and domain concepts. Annotators can then choose the appropriate values from this control vocabulary for the currently examined document, which can then be stored in the content model of the AH system.

The vocabulary used to describe the concepts for the experimental prototype were derived from a domain ontology, which described SQL. This ontology had been created a priori with the help of domain experts from the research group (see Figure 3-4).

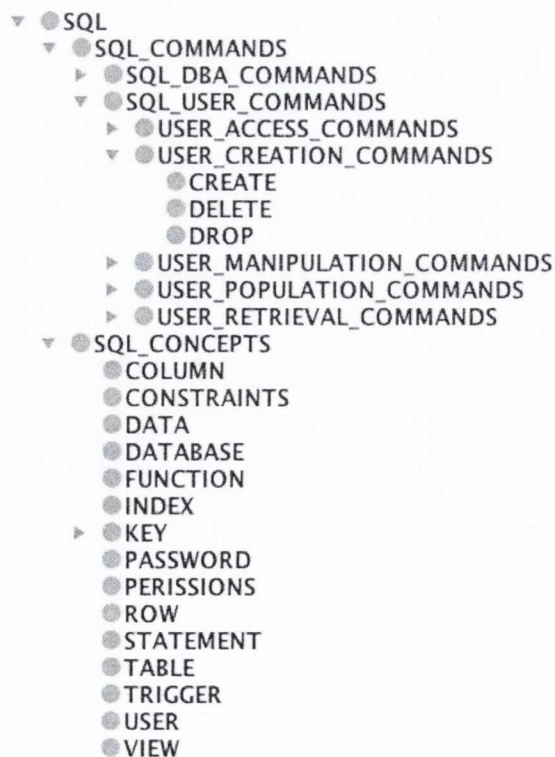


Figure 3-4. SQL ontology

3.4.2. Adaptation Process

As mentioned in section 3.1, the multimodel metadata-driven architecture has typically been applied to compose entire educational courses. However, the purpose of the system presented in this chapter lies in using the architecture to create educational information compositions that satisfy an informal user query. Therefore, the user query first needs to be matched to a Domain Model in order to identify the main concepts that cover the perceived information need. The implemented query interface allows users to compose a query from a set of domain-specific keywords (see Figure 3-5).

Step 1. Identify your question type:

What How

Step 2. Choose your intention:

Managing a Database Populating a Database Querying a Database Setting up a Database

Step 3. Select terms related to your question:

<input type="checkbox"/> ALTER	<input type="checkbox"/> CREATE TABLE	<input type="checkbox"/> DOMAIN	<input type="checkbox"/> LOAD	<input type="checkbox"/> REVOKE	
<input type="checkbox"/> ANALYZE	<input type="checkbox"/> CREATE TRIGGER	<input type="checkbox"/> DROP	<input type="checkbox"/> MERGE	<input type="checkbox"/> ROW	<input type="checkbox"/> TRUNCATE
<input type="checkbox"/> ASSERTION	<input type="checkbox"/> CREATE USER	<input type="checkbox"/> FOREIGN KEY	<input type="checkbox"/> NULL	<input type="checkbox"/> SCHEMA	<input type="checkbox"/> UPDATE
<input type="checkbox"/> BACKUP	<input type="checkbox"/> CREATE VIEW	<input type="checkbox"/> FUNCTION	<input type="checkbox"/> ORDER BY	<input type="checkbox"/> SELECT	<input type="checkbox"/> USE
<input type="checkbox"/> COLUMN	<input type="checkbox"/> DATA	<input type="checkbox"/> GRANT	<input type="checkbox"/> PASSWORD	<input type="checkbox"/> SET	<input type="checkbox"/> USER
<input type="checkbox"/> CONDITION	<input type="checkbox"/> DATABASE	<input type="checkbox"/> GROUP BY	<input type="checkbox"/> PRIMARY KEY	<input type="checkbox"/> SHOW	<input type="checkbox"/> VARIABLE
<input type="checkbox"/> CONSTRAINTS	<input type="checkbox"/> DATATYPE	<input type="checkbox"/> INDEX	<input type="checkbox"/> PRIVILEGE	<input type="checkbox"/> STATEMENT	<input type="checkbox"/> VIEW
<input type="checkbox"/> CREATE SCHEMA	<input type="checkbox"/> DELETE	<input type="checkbox"/> INSERT	<input type="checkbox"/> RENAME	<input type="checkbox"/> TABLE	<input type="checkbox"/> WHERE
<input type="checkbox"/> CREATE SEQUENCE	<input type="checkbox"/> DESCRIBE	<input type="checkbox"/> JOIN	<input type="checkbox"/> RESTORE	<input type="checkbox"/> TRIGGER	

Step 4. Submit!

Figure 3-5. Query Elicitation

Then, a domain-specific personal intent (goal) can be specified in order to provide the system with additional semantic information to compose an informed response. Additionally, a selected question type (what/how) indicates what type of response the user is hoping to receive. In the given example, this helps the system adapt to the user by either choosing more of an explanation-based or a more tutorial/example-based response. The system will generate a personalised response even if it only receives query keywords. However, including an intention and question type improves the results and presentation.

Following the query elicitation, the response composition process consists of three stages, *concept-level adaptation*, *content-level adaptation* and *presentation adaptation* (see Figure 3-6).

Stage 1: Concept-level Adaptation

First of all, the ad-hoc user query information is used in conjunction with the additional information held by the User Model in order to adapt the learning path across the Domain Model in a personalised manner. More specifically, the User Model and the Domain Model relationships are used in conjunction with the narrative in order to infer a personalised selection of the information space that should be presented to the user. This personalised selection depends on the domain knowledge of the user (held in the User Model), as well as the various relationships between domain concepts within the Domain Model. At the end of this stage, a Concept-Relationship Model has been created, which contains the selected concepts and relationships that best match the user’s personal information need.

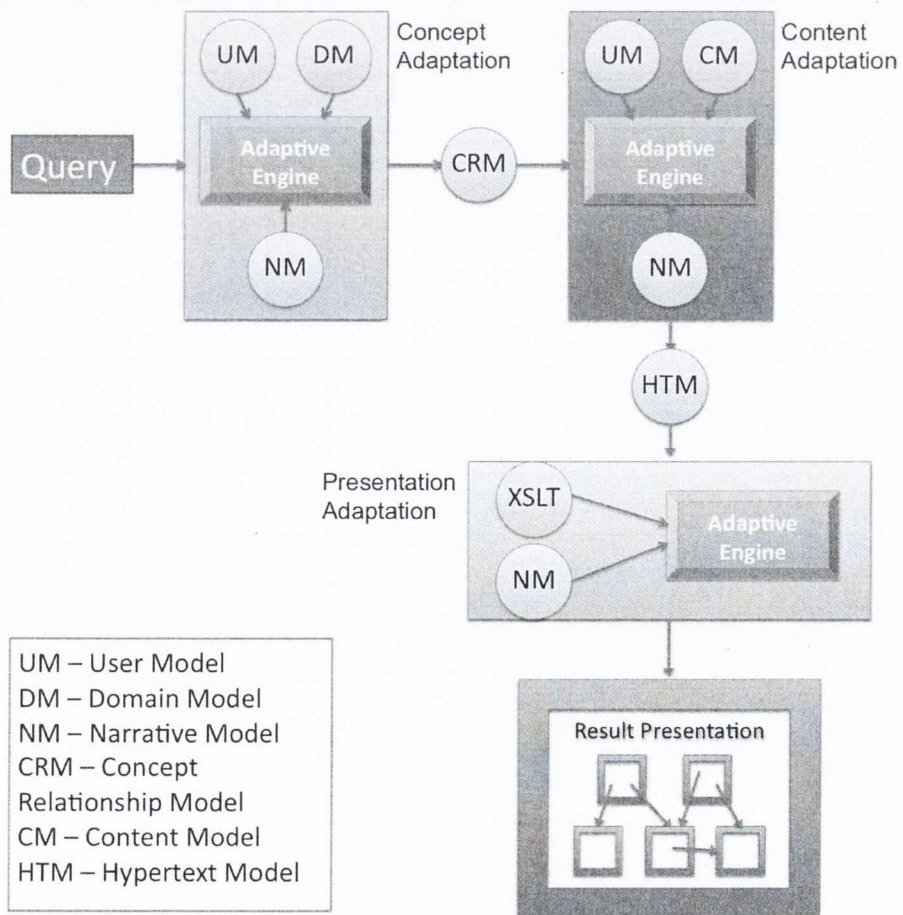


Figure 3-6. Composition generation process

Stage 2: Content-level Adaptation

In a second step, the strategy encoded in the narrative transforms the concept-relationship model into a coherent learning path. Again, using the User Model in conjunction with the Domain Model, the different concepts are sequenced in a form that is most suitable for a user. For example, if a user does not have sufficient knowledge of a particular concept, the prerequisites are sequenced before the actual concept. Additionally, for each concept, there exist different types of documents that serve varying educational purposes. By taking user preferences across these purposes, the result path can be personalised even further. For example, a user preferring examples to explanations will receive an appropriate result sequence according to these preferences. Then, for each of these “educational purpose/concept” pairs (e.g. introduction of SQL Select) an appropriate selection of documents is selected using the metadata that was generated in the document annotation phase. At the end of this second step, a complete Hypertext Model has been created, which contains a personally selected and sequenced portion of the information domain, as well as an adapted selection and sequence of documents with respect to the user preferences.

Stage 3: Presentation Adaptation

The final stage is concerned with transforming this Hypertext Model into a presentation that can be displayed to the querying user. Since the result model is encoded in a machine-processable XML format, the XSLT transformation capabilities are applied in order to produce a set of standard XHTML pages that can be viewed in a standard browser. The final result presentation to the user consists of a set of interlinked pages that display (i) a sequence of concepts to be visited (ii) a set of related concepts that will be beneficial to satisfy a user’s personal information need and (iii) a set of personally selected and sequenced documents that serve particular educational purposes for the selected concepts. The search results hence consist of an interlinked hypertext space that not only provides links to relevant documents, but it also guides and assists a user with a structured result composition. Figure 3-7 presents the overview page of selected concepts for the user query “PRIMARY KEY, GRANT” and Figure 3-8 presents a result page following the user selection of “Introduction of PRIVILEGE”. On this screen, users have a choice of up to three open-corpus documents that match the selected “educational purpose/concept” pair (e.g. Introduction of PRIVILEGE).

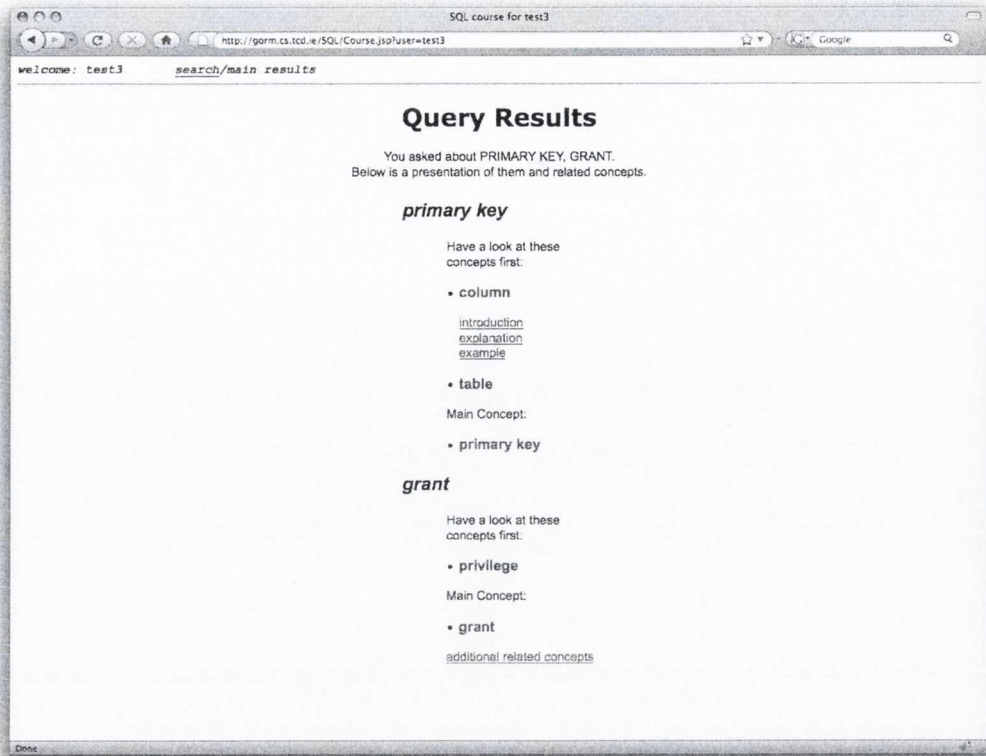


Figure 3-7. Result overview screen

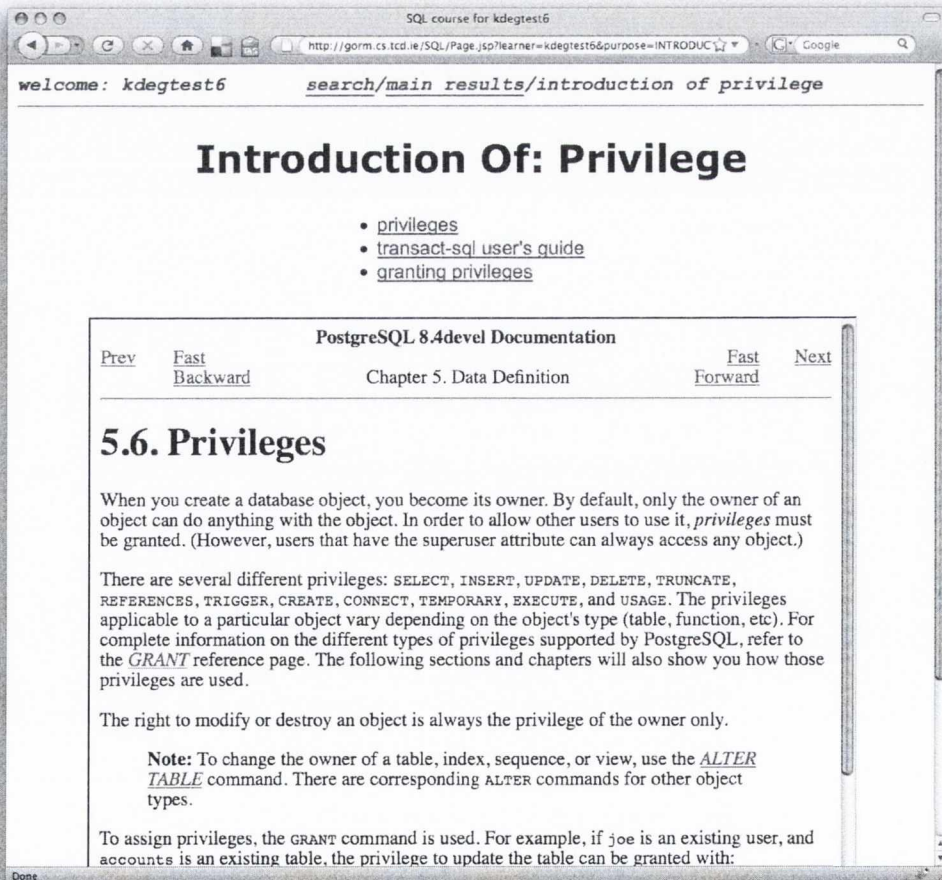


Figure 3-8. Content introducing the SQL Privilege concept

3.5. Evaluation

The presented Adaptive Composition prototype was applied in an authentic e-learning environment in order to evaluate the multimodel metadata-driven architecture for informal queries across open-corpus content. In order to carry out this evaluation, a case study was chosen on the area of teaching SQL from an open-corpus content base. For this purpose, a course test in the area of SQL was applied as the evaluation scenario. In particular, the prototype was evaluated in terms of:

- The system's metadata requirements
- The educational benefit (i.e. user effectiveness and efficiency)
- The usability from the students' perspective (i.e. user satisfaction).

3.5.1. Educational Benefit and User Satisfaction Hypotheses

Educational Benefit (user effectiveness and efficiency)

User effectiveness and efficiency typically refers to users being able to complete tasks successfully, quickly and with the least amount of effort. However, in the presented educational scenario, the user effectiveness and efficiency refers to the educational benefit that the system provides. To this end, one of the desired effects of an educational system is to motivate users to learn as much as possible on the given subject area (e.g. read more material, spend more time on the subject). This added motivation is also hoped to lead to higher task success rates and increased knowledge gains. The hypotheses regarding the educational benefit were therefore as follows:

- H1: Using adaptive compositions allows users to get higher scores in a course test than with conventional search systems.
- H2: Using adaptive compositions motivates users to explore and navigate across more content than using conventional search systems.
- H3: Users require less effort (in terms of number of queries) for finding relevant information with the adaptive composition system than using conventional search systems.

The corresponding null hypothesis for H1-H3 is that there are no differences between the adaptive composition system and conventional search systems.

User Satisfaction

User satisfaction is typically measured through usability questionnaires after completing given tasks with a system. The hypotheses regarding the user satisfaction for this study were as follows:

- H4: Users perceive the Adaptive Composition system to retrieve and present more relevant results than conventional search systems.
- H5: The Adaptive Composition system outperforms conventional search systems in terms of perceived usability.

Again, the corresponding null hypothesis for H4-H5 is that there are no differences between the adaptive composition system and conventional search systems.

3.5.2. Comparison to baselines

In order to test the above hypotheses (H1-H5), two additional prototypes were developed (referred to as IR+RF and IR+AH prototypes respectively), which applied personalisation across the same content base and using an identical query interface to allow a fair comparison with the Adaptive Composition system. However, both baseline systems used the conventional search system output of a ranked list. The system architectures of these two baseline systems are briefly described below.

IR+RF System Architecture

Traditionally, Information Retrieval systems operate over an inverted index that is created for the document base. In the Information Retrieval + Relevance Feedback (IR+RF) prototype, the harvested content cache has been indexed using standard indexing facilities of the Nutch²⁸ toolkit. The search facility used by Nutch uses a standard statistical term frequency algorithm. The search results are presented in a ranked list format to the users, similar to current web search systems such as Google.

In order to improve on this standard search system, Relevance Feedback functionality has been added to personalise the retrieval of results. The implementation of this functionality employs explicit user feedback, which requires users to indicate relevant results when they examine an initial result set. By subsequently clicking on “query refinement”, the user explicitly requests the system to expand the original query using

²⁸ <http://nutch.apache.org>

the most important terms from the documents that were marked as relevant (see Figure 3-9 and Figure 3-10).

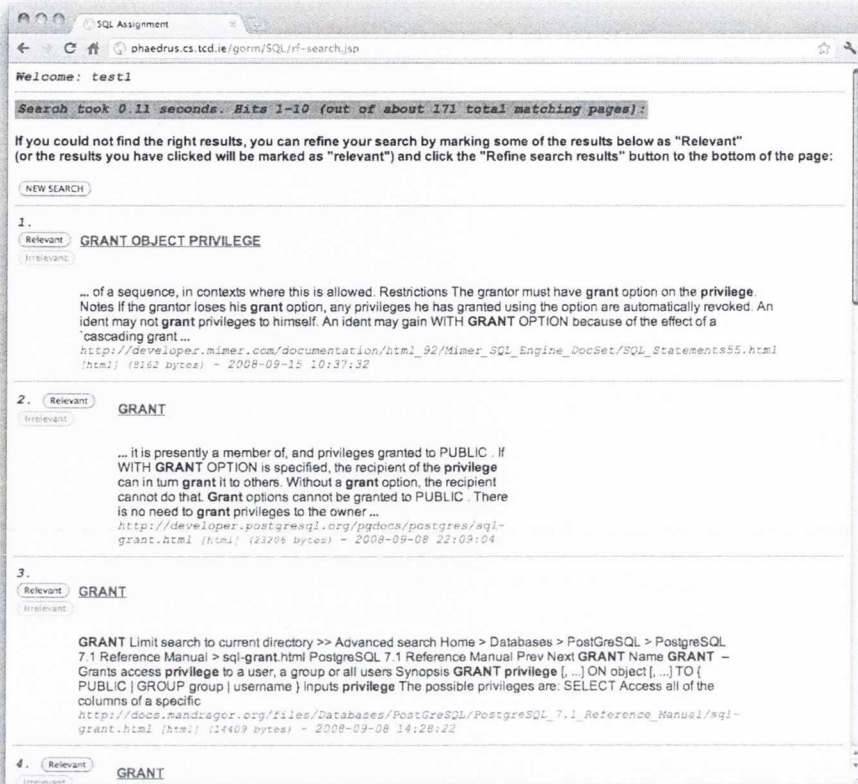


Figure 3-9. IR+RF prototype: top of result screen

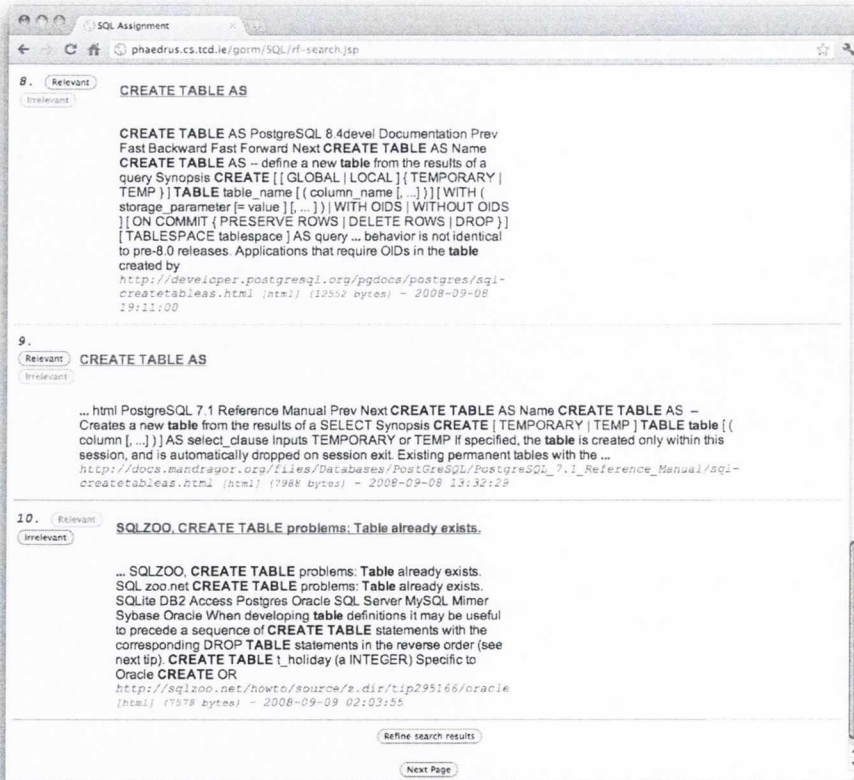


Figure 3-10. IR+RF prototype: bottom of screen, with query refinement button

IR+AH System Architecture

A third prototype combines components from both the Adaptive Composition system as well as the Information Retrieval system described above. More specifically, an Adaptive Hypermedia component is used in order to expand and personalise the original user queries (see Figure 3-11).

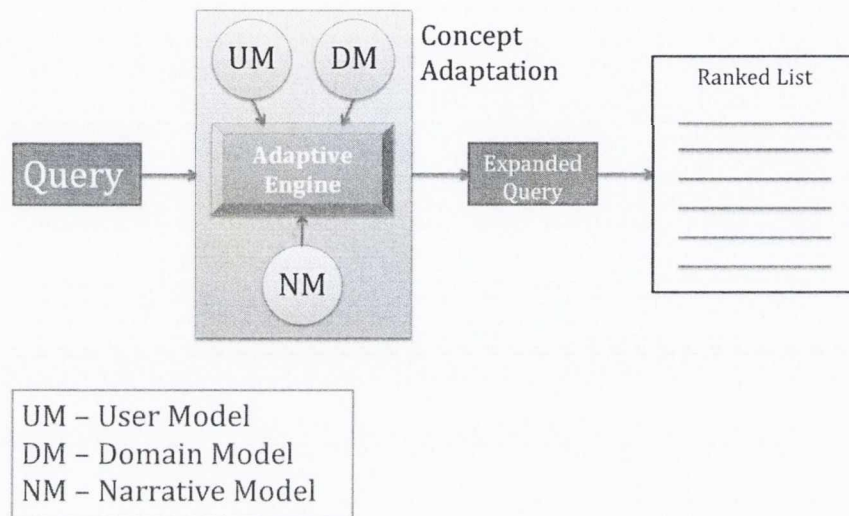


Figure 3-11. IR+AH prototype architecture

As explained in section 3.4.1, the response composition of the Adaptive Hypermedia system processes a user query in a first step in order to generate a personally relevant set of concepts in the form of a concept-relationship model (*Concept-level Adaptation*). For the hybrid AH-IR system, this exact first stage is used in order to create a personally relevant list of concepts. However, the relationships from the concept-relationship model are ignored, effectively turning the output into a bag of terms that can then be used in order to expand a traditional IR query. Similar to the Adaptive Hypermedia system, this set of concepts is generated using the Domain Model and the User Model in conjunction with the strategy encoded in the Narrative Model. However, as this set of retrieved concepts is only used to enhance a standard IR query, the ultimate output of this system is a ranked list identical to the IR+RF prototype.

3.5.3. Experimental Setup

The evaluation experiment consisted of 2 stages, namely (i) open-corpus content harvesting and crowd-sourced metadata generation and (ii) a task-based usage of the prototype systems by students in an authentic learning situation.

Content harvesting and Metadata Generation

During the first stage, a focused document cache was harvested using the OCCS, yielding approximately 15000 documents in the SQL subject domain. This cache was made available to researchers familiar with the domain using a web-based interface in order to view and annotate documents. Two categories of metadata were used to describe the content pages: the first category of metadata described the content of the document in terms of what were the primary and secondary concepts presented in the page, and which SQL commands these concepts described. As mentioned before, the vocabulary used to describe/identify the concepts and commands were derived from a domain ontology, which described SQL. This ontology had been created a priori with the help of domain experts from the research group. The second category of metadata described the document from the perspective of eLearning. This included an estimation of the prior knowledge required to understand a particular page and the type of document (e.g. tutorial, explanation, etc.).

Task-Based Experiment

During the second stage, a database course class of 35 students used the three prototypes as assistive tools for an authentic course test. Students were given 2 hours to complete all tasks and they were allowed to complete the tasks at their own pace.

Each student was given only one of the systems (IR+RF, IR+AH or the Adaptive Composition system) and the assignment of the course test tasks were based on a Latin square distribution. Each student received 3 out of a total of 6 tasks, which required the usage of their search system in order to gather and synthesise new knowledge. An example task would be *“What is a trigger? Explain how it can be used for automatically insuring integrity in a relational database. Give an example of a trigger command and explain how that example works.”*

The full task questions, as well as the usability questionnaires can be found in APPENDIX A.

The particular process for each student was as follows:

1. Each student was given a questionnaire in order to indicate their preferences as well as their *perceived* prior knowledge (see Figure 3-12). This information was used to generate the User Models.

SQL Assignment

phaedrus.cs.tcd.ie/gorm/SQL/Questionnaire.jsp

SQL Assignment Questionnaire

Section 1: Preferences

In this section you are asked to indicate your preferences in terms of learning material by choosing an option:

1. I find it easy to learn when presented with *Examples*:
 Strongly Agree Agree Disagree Strongly Disagree
2. I find it easy to learn when presented with *Exercises*:
 Strongly Agree Agree Disagree Strongly Disagree
3. I find it easy to learn when presented with *Explanations*:
 Strongly Agree Agree Disagree Strongly Disagree

Section 2: Knowledge and Experience

In this section you are asked about your knowledge of specific aspects of relational databases. For each topic, you can further refine your knowledge by assigning a number to each subtopic:

1. Do you understand the notion of a Database? No Yes
2. Do you understand the notion of a Table? No Yes
3. Do you understand the notion of Keys? No Yes
4. Do you know how to use SQL commands to create your own database? No Yes
5. Do you know how to use SQL commands to populate a database? No Yes
6. Do you know how to use SQL commands to query a database? No Yes
7. Do you know how to manipulate a database? No Yes
8. Do you know how to maintain a database? No Yes
9. Do you know how to manage the security of a database? No Yes

Figure 3-12. Pre-questionnaire

2. The users' *actual* prior knowledge was then gathered using a pretest, which consisted of specific questions about the domain of SQL. Each user was given a particular set of pretest questions that were similar to their allocated course test tasks (see Figure 3-13). This would allow an evaluation of the achieved knowledge gain after they completed their actual test.

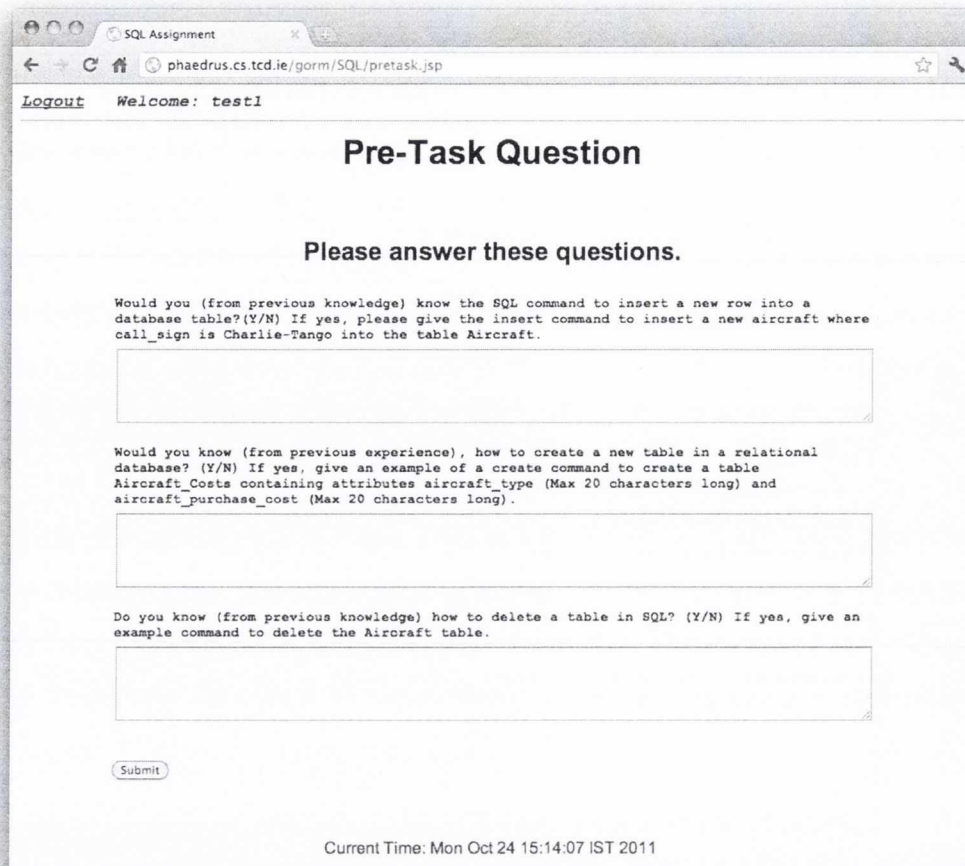


Figure 3-13. Pre-task questions

3. The course test tasks were then presented one at a time (see Figure 3-14), with students being able to use the tool they were assigned to in order to complete the task. As previously shown in figure 4, the query interface allowed students to indicate the type of task being performed (i.e. What/How), a query intention (e.g. Setting up a database), as well as one or more keywords to describe the question. The set of keywords was derived from the domain ontology describing the SQL subject domain.

During this phase, users' actions were tracked to identify particular trends in their search behaviour. The system collected the number of result pages viewed, as well as the time spent with the system in order to measure users' exploration motivation. Additionally, the query formulation was logged to identify the number of queries performed before answering a question, as well as the number of terms per query. Also, students' task answers were collected and corrected to measure the learning effectiveness of the system.

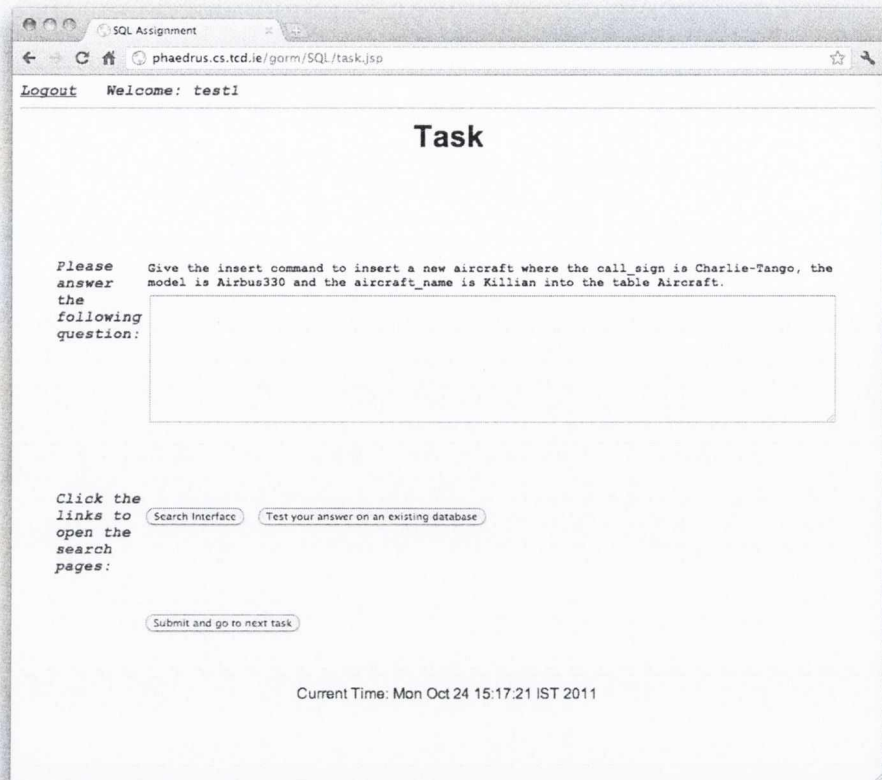


Figure 3-14. Task screen

4. After completing all three tasks, students were asked to complete an evaluation questionnaire, involving a series of standard usability questions (SUS) (Brooke, 1996). This independent questionnaire has been designed as “a reliable, low-cost usability scale that can be used for global assessments of systems usability”, regardless of the underlying application scenario and user base. SUS contains general usability questions such as “I thought the system was easy to use” or “I needed to learn a lot of things before I could get going with this system”. In addition to this more general usability questionnaire, users were asked to complete system-specific as well as free-text questions in order to express particular likes and dislikes.

3.5.4. Results

Content harvesting and Metadata Generation Results

A total of 20 annotators rated the cached documents according to the given metadata categories. 9,249 pages from the cache were annotated, of which 1,525 pages were rated as valid documents to be used by the prototypes.

The annotation tool recorded the start and end times of each annotation, but it was not possible to determine if, for example, the annotation tool was in the background while the user undertook another task. Because of this, a cut-off of 5 minutes was chosen in order to select valid annotations. This cut-off excluded less than two percent of the total annotations. An analysis of the annotation process revealed that the entire process took approximately 32 hours, with 90% of annotations being performed in less than 92 seconds (see Figure 3-15).

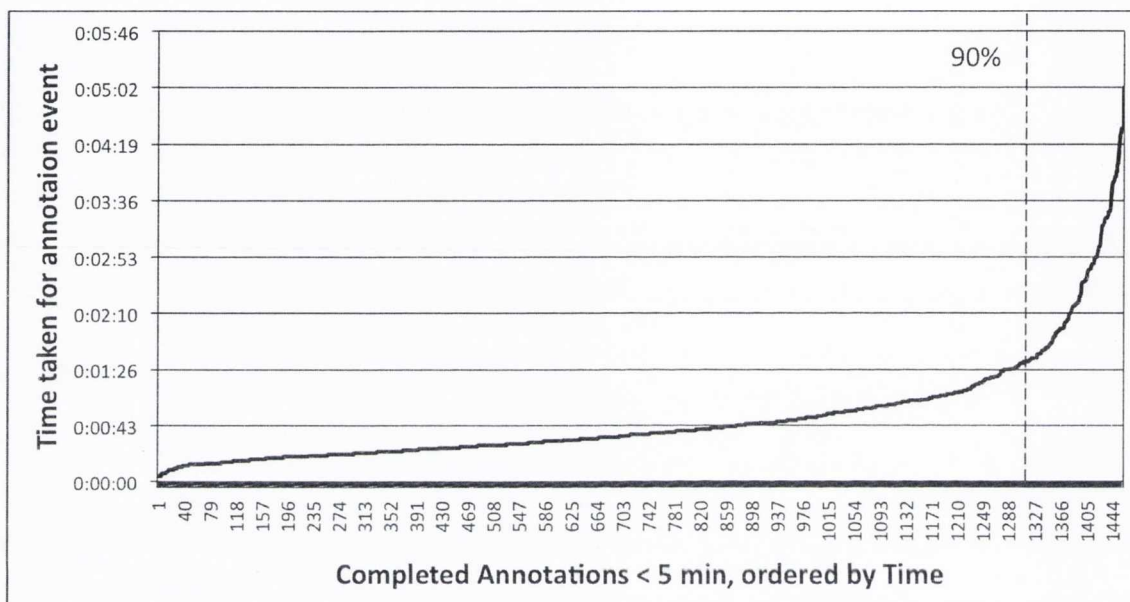


Figure 3-15. Duration of annotation events that produced complete descriptions in under five minutes

Task-Based Experiment Results – Educational Benefit

To quantify the students’ knowledge gain, the pretest scores were compared with the actual course test scores. This was calculated by scoring students’ pretest answers in the range from 0 to 5 (0 representing complete absence of knowledge to complete the pretest and 5 representing successful completion of the pretest) and scoring the course test itself in the range of 0 to 5 (0 representing complete failure and 5 representing complete success).

In terms of group balancing, the pretest scores revealed that the prior knowledge was similarly low for each system group. On average, students who later used the IR+RF system scored 0.27 out of 5 on their pretests, the IR+AH group 0.31 out of 5 and the Adaptive Composition system group 0.28 out of 5. These differences were not found to be statistically significant in ANOVA tests ($p=0.583$). It could therefore be assumed

that the groups were equally balanced in terms of prior experience with the subject domain.

The knowledge gain results revealed that there was a marginal difference between the different user groups. On average, the IR+RF group scored 3.82 out of 5, the AH+IR group 3.84 out of 5 and the Adaptive Composition system group 4.25 out of 5. ANOVA tests revealed that these results were not statistically significant ($p=0.082$). Similar results were found when comparing the systems across each of the 6 different tasks. However, a Tukey post-hoc test showed that there were much lower significance level values for the comparison of the Adaptive Composition system to the IR-based systems ($p=0.068$ with IR+RF and 0.161 with IR+AH) than the IR-based systems compared directly ($p=0.998$). Although the overall differences are not significant at the rigorous significance cut-off value of 0.05, this clear trend may be an indication that the Adaptive Composition system helped students in gaining more knowledge, hence providing partial support for hypothesis *H1*.

In terms of user motivation (to explore and navigate across more content), the analysis of the students' result page viewing behaviour revealed significant differences between the Adaptive Composition system and the two IR baselines. On average, users of the Adaptive Composition system looked at 4.30 documents per query, whereas users of the IR based systems only looked at 1.81 (IR+RF) and 1.78 (IR+AH) documents respectively ($p=0$). This finding was observed across each of the 6 tasks (see Figure 3-16).

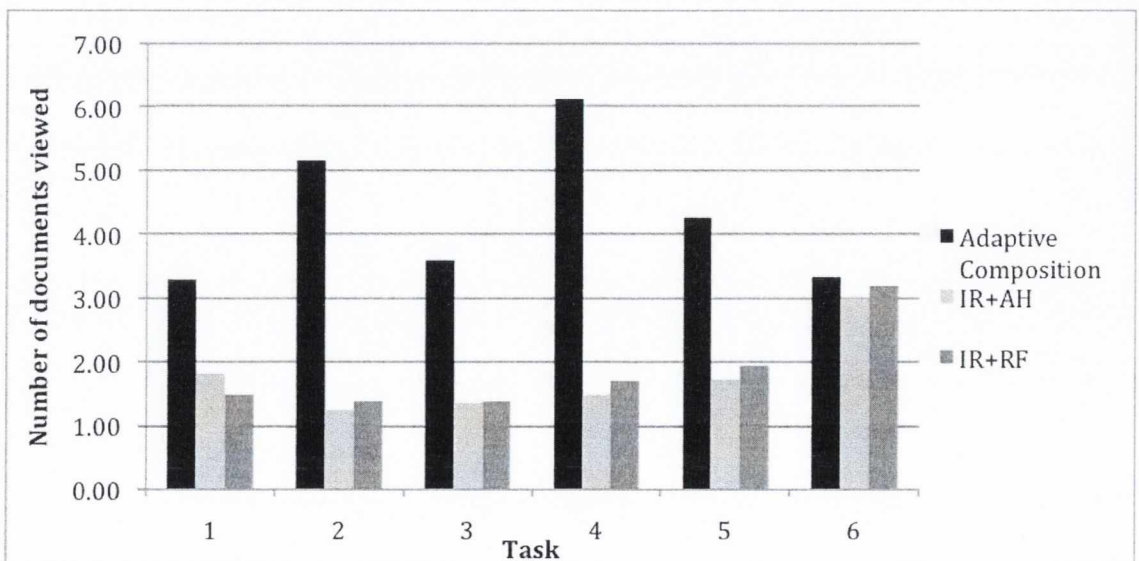


Figure 3-16. Number of documents that users looked at per query

In total, users of the Adaptive Composition system looked at 280 documents compared to 63 (IR+RF) and 72 (IR+AH) documents respectively. This also led to a significant difference in terms of overall time spent with the system: 32:12 minutes for the Adaptive composition system, 25:40 for the IR+RF system and 24:06 minutes for the IR+AH system ($p=0.04$). These findings clearly point towards the fact that users felt encouraged to explore the result space composed by the Adaptive Composition system (e.g. introductions, explanations, examples, related concepts). As the application of the system was educational in nature, it could therefore be argued that it provided the educational benefit of motivating students, providing clear support for hypothesis **H2**. Although these findings could also lead to an alternative conclusion that the system provided many irrelevant results, the answers given by the students in their questionnaires revealed that this was not the case (presented in the user satisfaction results below).

The analysis of the students' query behaviour revealed that there was no statistically significant difference in terms of user effort for finding relevant information. On average, users of the Adaptive Composition system issued 2 queries per task, IR+RF users 1.7 queries and IR+AH users 1.9 queries ($p=0.67$). These results show that users had to issue a similar amount of queries in order to find relevant information, therefore not providing support for hypothesis **H3**. However, when coupling this finding with the increased page view count of the Adaptive Composition system, it may be argued that students were more effective in terms of viewing more documents with the same amount of queries, hence providing at least partial support for hypothesis **H3**.

Task-Based Experiment Results – User Satisfaction

As mentioned in the experimental setup section (section 3.5.2), users were asked to complete a set of usability questionnaires after completing the course test. The first set of questions required users to agree with a set of statements on a Likert scale from 1 (completely disagree) to 5 (completely agree), including the Standard Usability Scale (SUS). A second set of free text questions also allowed users to express any particular likes and dislikes of their given prototype system.

When asked to agree or disagree with the statement “*I found the search system returned relevant search results for my query*”, there were no statistically significant differences between the systems. The Adaptive Composition system scored an average of 3.91, the IR+RF system 3.73 and the IR+AH system 3.58 ($p=0.59$). Similar results were also

found when asked about “*I found the search system returned irrelevant search results for my query*”. The Adaptive Composition system scored an average of 2.58, the IR+RF system 2.54 and the IR+AH system 2.50 ($p=0.98$). These results both indicate that there was no significant difference between the prototypes in terms of returning relevant results to users. However, as mentioned in the educational benefit results above, this could be interpreted in a positive light for the Adaptive Composition system. Since users chose to view a greater number of results, these findings confirm that the additional page views were not due to irrelevant results, but rather due to the motivational effect of the result composition.

When asked to agree or disagree with the statement “*I found the presentation of the search results helpful*”, there were again no statistically significant differences between the systems. The Adaptive Composition system scored an average of 3.67, the IR+RF system 3.27 and the IR+AH system 3.25 ($p=0.22$). Similarly, the SUS scores provided no statistically significant differences between systems, 60.20 for the Adaptive Composition system, 62.5 for the IR+RF system and 72.27 for the IR+AH system ($p=0.153$).

Overall, since these questionnaire answers provide no statistically significant differences between the three prototypes, hypotheses **H4** and **H5** can therefore not be directly supported. However, when analysing the free text answers of the students, it is possible to find more specific likes and dislikes of users.

When asked “*What did you like most about the search system?*” 6 out of the 12 Adaptive Composition users mentioned the relevance of the results and 4 students provided positive comments regarding the result presentation. Moreover, 4 out of the 12 students mentioned that they particularly appreciated the result compositions, saying for example that they liked “the different sections within the results” or the “optional resources”. On the other hand, when asked “*What did you like least about the system*” 5 out of the 12 students mentioned the interface design of the prototype. This might provide some indication as to why the Adaptive Composition prototype scored relatively low in terms of usability from the students’ perspective. The interface design was very basic and might therefore not have been aesthetically pleasing. Especially considering that the interface was relatively novel compared to the established Information Retrieval interface paradigm, it would have been very important to make the composition layout both functional and more visually attractive.

For the IR-based systems, there were no particular trends in terms of likes or dislikes. However, when analysing answers across all prototypes, 9 out of the 35 students mentioned their dislike of having to choose terms from a restrictive list for eliciting queries.

3.5.5. Discussion

The evaluation results have provided initial evidence of the potential benefits and limitations of applying the multimodel metadata-driven architecture for informal queries across open-corpus content. The case study on the area of teaching SQL has provided an authentic evaluation environment and has revealed several benefits of the compositional approach. First of all, results have shown that users are encouraged and motivated to explore and navigate across more educational content without having to issue more queries. This finding is in line with other adaptive educational systems, which have for example reported increased user motivation when providing adaptive hints for course tests (Hsiao, 2009). Moreover, it is confirmed by the student questionnaire answers that these increased document views are not a result of irrelevant results being returned by the system.

However, from a perceived user satisfaction perspective no statistically significant differences were observed compared to conventional search systems. This could potentially be attributed to the fact that the Adaptive Composition system had a very basic interface and that it might not have been as intuitive to use compared to the well-established Information Retrieval interfaces. Another possible explanation for the lack of statistical significance in the results might stem from the fact that the number of participants (35) was rather low, particularly since each user experienced only one of the systems. A comparative evaluation, where each user experiences each of the systems might provide more of an indication for the comparative user satisfaction. Similarly, although no statistically significant differences could be found regarding student knowledge gain, this may be attributed to a relatively high ability of students, as well as a low degree of difficulty of the actual tasks (almost all students answered the tasks correctly). It is therefore possible that a similar study with harder tasks may better highlight the additional guidance of the compositional approach.

Overall, the evaluation of the architecture in the educational case study has proven that the previously closed-corpus design principles can be applied successfully for informal

queries across completely externally-sourced open-web documents without any penalty in terms of efficiency, effectiveness and satisfaction. However, the fact that users had to choose from a set of predetermined keywords for query elicitation has been identified as a limitation of the architecture and needs to be addressed in future revisions. Moreover, the content harvesting and crowd-sourced annotation results have shown that a significant amount of manual effort is involved in generating the necessary metadata for the content cache (around 32 hours for annotating 1525 relevant pages). Although such an annotated cache might not necessarily have to be updated regularly for teaching established concepts, there are many other (non-educational) domains where information is constantly changing and therefore requiring a more responsive process for incorporating new documents.

3.6. Conclusions

This chapter has described an initial adaptive compositional approach, which applies multi-model Adaptive Hypermedia design principles across completely externally-sourced open-corpus information. The presented architecture adaptively loads and applies adaptation strategies (narratives) across a multitude of models, including domain, user and content models. The implementation has been shown to consist of a modular, component-based framework that allows a variety of model manipulations, including extensions for semantic querying and reasoning.

In order to evaluate this architecture for satisfying informal queries across open-corpus content, an e-learning prototype has been applied in the area of teaching SQL. Results of this evaluation have shown that the compositional approach to result delivery motivates users to explore more resources while issuing the same number of queries. Coupled with the fact that students were also able to achieve a sizable knowledge gain, it has been shown that the approach can successfully provide educational benefit to users. In terms of user satisfaction, it has been shown that there is partial evidence for the potential benefits of the compositional approach, most notably the fact that students expressed their particular liking towards the result relevance, presentation and composition.

To facilitate the initial integration of open-corpus content into the metadata-driven architecture, a crowd-sourced approach has been applied for generating the necessary

metadata for adaptation and composition. This approach has shown that the metadata generation step can be run independently of the actual content creation while still maintaining sufficient quality to provide relevant results to end users (given the metadata annotators have sufficient domain knowledge to produce accurate annotations). Moreover, for the educational case study shown in this chapter, such an annotated content cache arguably requires infrequent updating in order to stay relevant.

However, there are many other (non-educational) domains where information is constantly changing and which require open-corpus content to be integrated at a quicker and less labour-intensive rate. In such cases, it might be impractical to apply the annotation process shown in this initial study. However, while the metadata requirements are relatively strict for educational applications (e.g. requiring accurate descriptions of the educational purpose of a document), this is not necessarily the case for other domains. The metadata-driven architecture of the Adaptive Composition system should therefore be extended in order to integrate both metadata-rich as well as metadata-sparse information. To this end, additional open-corpus extensions need to be integrated to include unstructured open-corpus content. This integration should preserve the multi-model adaptation design principles contained in the initial architecture in order to maintain the ability to apply multiple levels of adaptation for result composition. In particular, by coupling this adaptivity across metadata-rich closed-corpus content and unstructured open-corpus content (i.e. heterogeneous information), the extended architecture should be able to provide additional usability benefits to users across other (non-educational) domains.

A second limitation of the initial Adaptive Composition architecture has been identified in terms of its query elicitation capabilities. In order to specify a query, the initial prototype requires users to pick terms from a set of domain-specific keywords. In the evaluation results, it has been shown that users do not appreciate this type of query specification, particularly considering the fact that free text search can be found across all modern web search engines. Moreover, for bigger domains it would be infeasible to display all domain keywords in this manner. Also, if metadata-sparse information is to be handled and integrated into the system, the architecture needs to be extended in order to query across large, unstructured document collections.

In conclusion, this chapter has provided initial evidence for the benefits of providing adaptive information compositions across open-corpus information. However, several

limitations in terms of open-corpus content capabilities have been identified, which need to be addressed in order to apply the architecture across large, dynamic and heterogeneous content bases. In the following chapter (chapter 4), an extended architecture is presented, which aims to overcome these limitations in order to provide adaptive retrieval and composition of heterogeneous information sources.

4 ARCHING - Adaptive Retrieval and Composition of Heterogeneous Information sources for personalised hypertext Generation

4.1. Introduction

Chapter 3 has presented an initial iteration of an adaptive compositional approach to information retrieval and delivery using a state-of-the-art Adaptive Hypermedia architecture. This architecture has been shown to successfully generate adaptive compositions across externally-sourced open-corpus information. However, it has also been shown that the process of integrating such information has been limited by the fact that the architecture requires extensive amounts of metadata about the content corpus. Moreover, the initial evaluation has revealed limitations in terms of allowing users to freely specify keyword queries in a manner that is commonly found in modern search systems.

In order to overcome these limitations, the initial approach therefore needs to be extended in terms of flexibility for integrating both metadata-rich closed-corpus information as well as metadata-sparse open-corpus information. Such additional (lightweight) open-corpus content handling and adaptation functionalities are required in order to apply the adaptive compositional approach to domains where information is constantly changing on the open web.

This chapter first presents the design principles for an extended architecture of the Adaptive Composition system in order to provide the required functionalities (section 4.2). These principles are both influenced by the findings of the initial Adaptive

Composition prototype presented in chapter 3, as well as the state of the art in AH and PIR presented in chapter 2.

Section 4.3 presents an extended architecture called ARCHING (Adaptive Retrieval and Composition of Heterogeneous INformation sources for personalised hypertext Generation), which integrates lightweight query and retrieval adaptation functionalities, while fully retaining the adaptive composition and presentation capabilities of the first iteration. This architecture allows a novel compositional approach to information retrieval and delivery, which can generate adaptive information compositions from closed-corpus and open-corpus information sources according to multiple dimensions of adaptation.

A prototype implementation of ARCHING is presented in a customer care scenario (section 4.4), which provides authentic information needs, heterogeneous data sources, as well as real-life evaluation possibilities. The prototype is evaluated in terms of benefits to users compared to a purpose-built, non-adaptive search system (section 4.5). Results from this task-based evaluation show that the compositional approach significantly enhances a user's efficiency, effectiveness and satisfaction.

Finally, section 4.6 concludes this chapter with a discussion of the overall findings, as well as alternative composition possibilities.

4.2. Design Principles

The evaluation of the initial Adaptive Composition system has provided initial evidence for the benefits of adaptive compositions across open-corpus information. In particular, it has been shown that the notion of creating adaptive information compositions, presentations and navigations can provide benefits in terms of user effectiveness and motivation.

The first design principle for an extended composition system therefore lies in *retaining the adaptive composition, presentation and navigation capabilities* of the initial architecture. Moreover, the compositions should be generated according to *multiple user dimensions*, such as a user's knowledge level, query intent or task context. In order to retain these functionalities, the extended architecture should therefore build on the *modular, multimodel-driven approach* provided by the first

composition system. In addition to this, the composition system architecture should retain lightweight modelling for determining user properties (e.g. query intent), as well as allowing more long-term user properties to be held in the user model, e.g. user knowledge, language preferences, device capabilities.

However, several additional design principles are necessary in order to open up the initial multimodel metadata-driven approach. First of all, the extended architecture needs to be able to *operate over metadata-rich closed-corpus as well as metadata-sparse open-corpus information*. The architecture should allow this *heterogeneous, multilingual content to be retrieved and composed in a fully integrated way*, enabling users to benefit from the rich markup and adaptation possibilities of closed-corpus information as well as the quantity and breadth of multilingual open-web information. This open-corpus content should be adaptively retrieved and integrated on-the-fly and without requiring structural or descriptive metadata. Moreover, the architecture should allow users to specify their initial information needs using conventional *free-text keyword queries*.

In order to integrate these functionalities, it is therefore necessary to add open-corpus techniques and technologies commonly found in Personalised Information Retrieval (PIR) systems. In particular, *keyword-based query, retrieval and classification capabilities* can be employed in order to adaptively retrieve and classify information from the open web. This should include techniques for the adaptation of the initial user query, the selection of the retrieval engine and content source as well as the translation and classification of open-web results. As outlined in section 2.5.3, it is crucial to *flexibly integrate the PIR techniques with the adaptive composition, presentation and navigation techniques* in order to align them into an overarching strategy. Through this integration, it is possible to make full use of the complementary benefits of AH and PIR techniques across both closed-corpus as well as open-corpus adaptation techniques.

Overall, the design principles outlined above should assure that the benefits of adaptive compositions, navigations and presentations are retained, while allowing system prototypes to satisfy free-text keyword queries and to flexibly integrate metadata-sparse open-corpus content. This combination of principles and techniques are aimed at providing the type of adaptive information compositions outlined in the research question (section 1.2) to enhance a user's *effectiveness, efficiency and satisfaction*.

4.3. Architecture

As outlined in the design principles above, the extended architecture (called ARCHING) should preserve the adaptive composition, presentation and navigation principles of the first iteration. Moreover, the modular, multi-model driven approach should be retained in order to provide adaptation according to multiple dimensions of adaptation.

Section 4.3.1 gives a high-level overview of the extended architecture, which fully preserves the strategy interpretation and model manipulation capabilities of the initial architecture, while adding open-corpus manipulation capabilities. Section 4.3.2 presents the technological implementation of the additional components, which utilise statistical keyword-based adaptation techniques for open-corpus content manipulation.

4.3.1. Architecture Components & Capabilities

Figure 4-1 provides a high-level overview of the various components and capabilities of the ARCHING system.

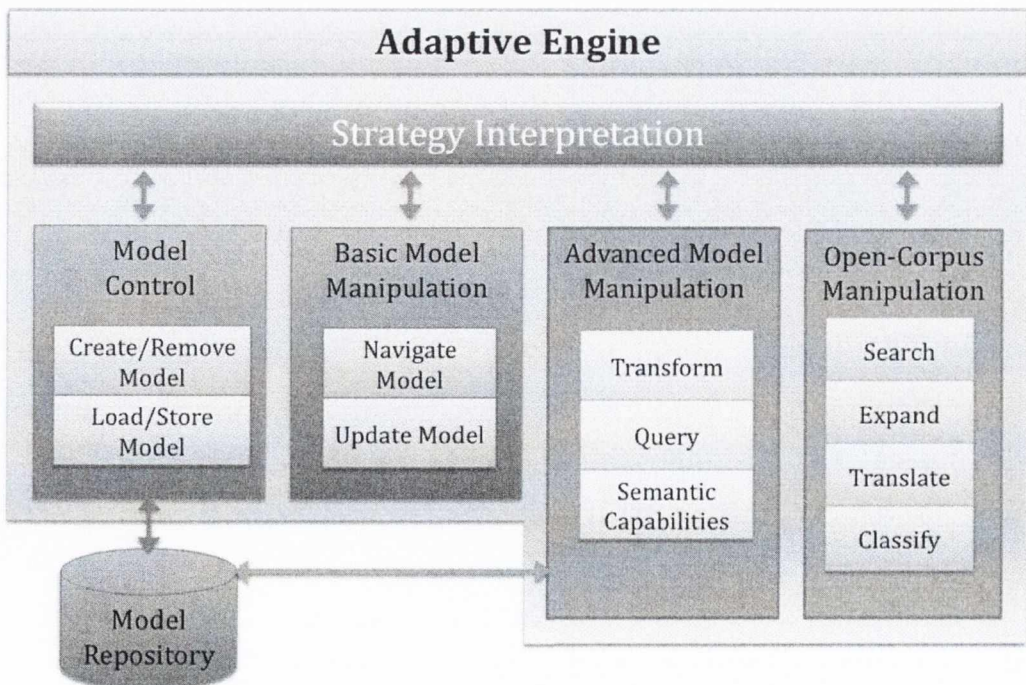


Figure 4-1. ARCHING Component Architecture

As shown in this diagram, ARCHING extends the original multimodel metadata-driven architecture with open-corpus manipulation capabilities. These capabilities include *Open-Corpus Search*, *Query Expansion*, *Text Translation* and *Document Classification*. Moreover, additional design flexibility has been integrated in order to perform open-corpus manipulation either directly through the Adaptive Engine, or through asynchronous requests.

4.3.1.1. Search

First of all, the integration of search components is aimed at allowing users to freely specify keyword queries for indicating their information need. It is also fully compatible with the closed-corpus content model and allows keyword searches across entire documents. This alleviates the need for selecting keywords from a narrow list of terms and brings the user experience closer to conventional search engine query elicitation.

Secondly, the components allow searches to be performed across a variety of information sources, including (i) closed-corpus content and large open-corpus content caches and (ii) the open-web and specific subdomains of the open-web. Results retrieved from each of these information sources are handled identically to the models held by the Adaptive Engine, allowing a full integration of heterogeneous results into adaptive information compositions and presentations. Search selections and sequences are specified in the narrative by defining search parameters in parameter models. The range of search engines and parameters that are currently supported will be discussed in the technological architecture (section 4.3.2.1).

4.3.1.2. Expand

As shown in chapter 2, one of the most popular ways to perform personalised search is through query adaptation. In addition to the search engine and source selection parameterisation mentioned above, ARCHING contains a component that generates expanded queries based on (i) the original user query and (ii) textual content that reflects the user's current information interest.

Similar to the added search components, the query expansion can be called directly from the narrative in order to fit into an overarching strategy. For example, in a certain

context it might be desirable to issue the original query to a particular search engine/information source, whereas in other contexts the query needs to be expanded for result refinement/diversification.

Since ARCHING contains a variety of search options (in terms of search engines and information sources), the query expansion module needs to be able to support varying query requirements. This is particularly important in terms of the query input requirements of the underlying search engine (e.g. stemming or term boosting). The technological details of these requirements along with the query expansion algorithm will be explained in section 4.3.2.2.

4.3.1.3. Translate

In addition to the adaptation dimensions addressed in the initial composition system (e.g. user knowledge, query intent), ARCHING has been designed to provide adaptive information compositions from *multilingual information sources*. To this end, text translation capabilities have been integrated into the architecture in order to translate (i) a user's query and/or (ii) retrieved information.

The ability to adaptively retrieve and compose multilingual information presents novel opportunities for increasing information diversity, as well as for addressing information sparsity in a particular language. Moreover, as will be shown in chapter 6, there are a number of possibilities for supporting *multilingual users* across related content from multilingual sources.

4.3.1.4. Classify

As presented in chapter 2, many personalised search systems make use of categorisation and classification in order to map search results to underlying directories (e.g. ODP categories). Such capabilities enable the categorisation of unstructured open-web results into more semantically rich knowledge structures.

The design of ARCHING contains such classification capabilities, which allow the mapping of unstructured open-corpus documents to ontological concepts from the domain model. These capabilities make use of statistical keyword-based algorithms for training models, which can then be used at runtime for on-the-fly classifications.

The design is again fully integrated with the strategy interpretation and model manipulation capabilities, allowing narratives to sequence the classification at appropriate stages. For example, in the case of a user query not being covered by the closed-corpus content base, open-corpus results can be retrieved and classified in order to determine the conceptual scope of the query. Following this, structured queries regarding these concepts could complement the open-corpus results with related closed-corpus information. The technological details of the current implementation algorithms for these capabilities will be explained in section 4.3.2.4.

4.3.1.5. Additional Design Flexibility

In addition to the presented architectural extensions, ARCHING allows increased flexibility for using these components. In particular, the additional flexibility is aimed at prototypes that need to be able to cope with delays introduced by open-web queries.

The ARCHING design therefore allows the specification of asynchronous requests, whereby the narrative only specifies the query parameters without running the actual search. This ensures that the result composition can be returned quickly to the user without requiring all search requests to be completed. The open-search result retrieval can then be performed using asynchronous client-side technologies, hence enabling a more responsive user experience.

4.3.2. Technological Architecture

The technological architecture builds on the modular framework presented in the initial composition system and provides fully integrated components for the functionality described above.

As shown in Figure 4-2, the open-corpus modules are integrated in the Adaptive Engine and can be accessed identically to the original manipulation components, allowing narratives to compose both closed-corpus and open-corpus results. The implementations of the *Search*, *Expand*, *Translate* and *Classify* modules make use of open-source libraries, as well as RESTful web services. These libraries and services have been wrapped using custom Java modules in order to fully integrate them into the Adaptive Engine architecture. Moreover, in order to provide the *Additional Design*

Flexibility mentioned in section 4.3.1, client-side scripting technologies have been added to enable asynchronous search requests.

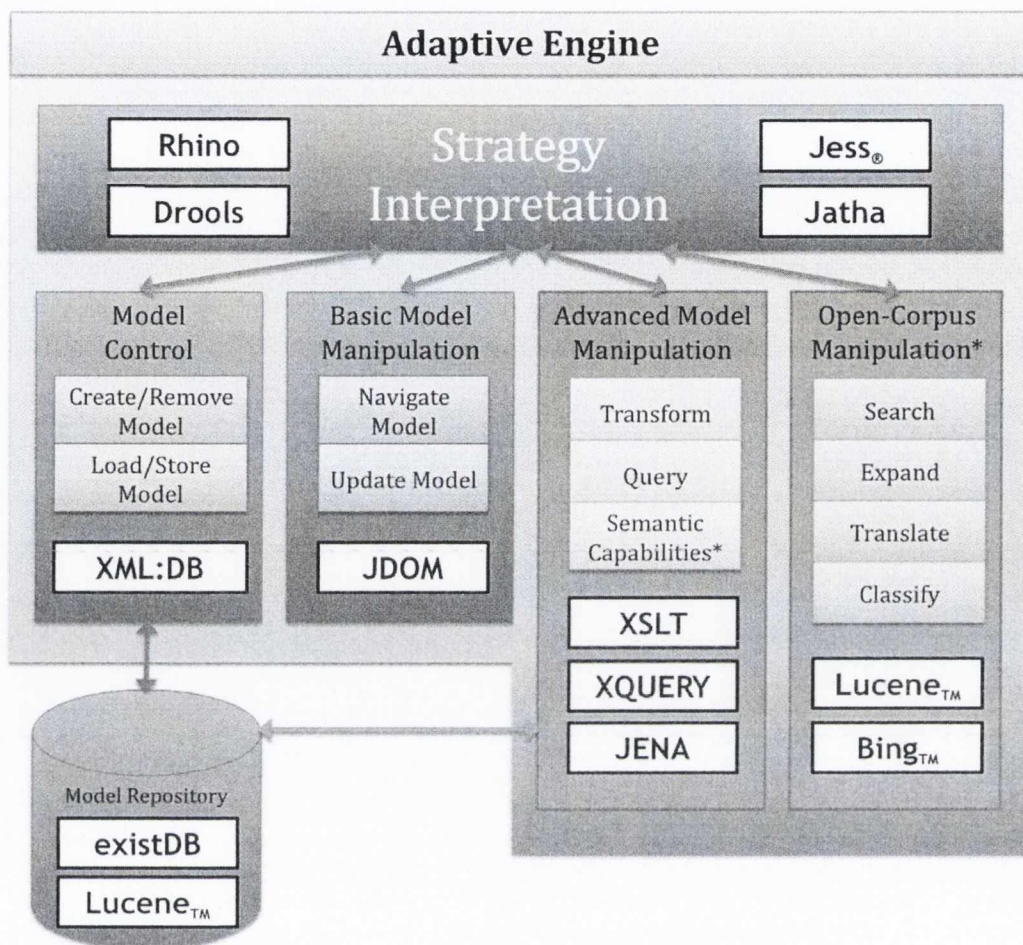


Figure 4-2. ARCHING Technological Architecture

4.3.2.1. Search

ARCHING integrates search functionalities that operate over (i) structured closed-corpus and unstructured open-corpus content caches and (ii) the open-web.

Structured Closed-Corpus / Unstructured Open-Corpus cache search

The search functionality over closed-corpus and open-corpus content caches has been integrated using *Lucene*²⁹, an open-source information retrieval software library. This library provides a range of modules for content indexing, analysis and keyword search.

The analysis modules used by the implementation in ARCHING are the *SnowballAnalyzer*³⁰ for English, *GermanAnalyzer*³¹ for German and *FrenchAnalyzer*³²

²⁹ <http://lucene.apache.org/>

for French. These modules perform basic stop-word removal, as well as word stemming.

The document scoring module in Lucene calculates relevance scores for the retrieved documents with respect to the given query. The scoring function utilises *term frequency-inverse document frequency (tf-idf)* style term weighting and *Vector Space Model (VSM)* query-document similarity rating³³. *Term frequency tf* (for term *t* in document *d*) correlates to the term's frequency, defined as the number of times term *t* appears in the currently scored document *d*. Documents that have more occurrences of a given term hence receive a higher score. *Inverse Document Frequency idf(t)* correlates to the inverse of the number of documents in which the term *t* appears. This means rarer terms are given higher contribution to the total score. The VSM score of document *d* for query *q* is the *cosine similarity* of the weighted query vectors *V(q)* and *V(d)*:

$$\text{cosine-similarity}(q,d) = \frac{V(q) \cdot V(d)}{|V(q)| |V(d)|}$$

These modules have been integrated to allow the execution of freetext queries across (i) structured closed-corpus content as well as (ii) unstructured open-corpus content caches.

First of all, in order to provide query functionalities across structured closed-corpus content, the *existDB* database has been configured to incorporate *Lucene-based modules*³⁴ for analysis, indexing and keyword search. This allows the execution of fulltext XQuery/XPath queries across the complete textual content of XML documents. Coupled with the structured retrieval capabilities of XPath/XQuery, narratives can therefore combine metadata-based searches with fulltext scoring functionalities.

Additionally, custom Java modules have been integrated in order to run Lucene-based freetext queries independently of the *existDB* database. This allows searches across self-contained, large-scale content cache indices without requiring the content to be held in a structured database.

³⁰ http://lucene.apache.org/java/2_9_1/api/contrib-snowball/org/apache/lucene/analysis/snowball/SnowballAnalyzer.html

³¹ http://lucene.apache.org/java/2_9_1/api/all/org/apache/lucene/analysis/de/GermanAnalyzer.html

³² http://lucene.apache.org/java/2_9_1/api/all/org/apache/lucene/analysis/fr/FrenchAnalyzer.html

³³ http://lucene.apache.org/java/2_9_1/api/core/org/apache/lucene/search/Similarity.html

³⁴ <http://exist-db.org/lucene.html>

There are several parameters that can be set for searching with this custom module:

- *Query*: the keyword query to be used for searching across the index
- *Number of results*: the number of results to be returned
- *Page*: the starting point for the search, e.g. if the number of results is set to 10 and the page number is set to 2, results 11-20 will be returned
- *Index location*: physical base location of the search index
- *Index name*: name of the search index (excluding language parameter)
- *Language*: the natural language of the query, used for determining the appropriate index and analysis module

Open-Web search

In addition to supporting searches over structured and unstructured content caches, the ARCHING implementation contains open-web search functionalities through the Bing API³⁵. Custom Java modules have been implemented to fully integrate this RESTful service into the ARCHING architecture.

As with the Lucene module, there are several parameters that can be set for customising searches:

- *Query*: the keyword query to be used for Bing search
- *Number of results*: the number of results to be returned
- *Offset*: the starting point for the search, e.g. if the number of results is set to 10 and the offset is set to 11, results 11-20 will be returned
- *Site*: the domain to be searched across, e.g. "tcd.ie". If this parameter is left blank, the search will be performed across the open web.
- *Source*: the type of content to be returned, e.g. web, video, images
- *Language*: the natural language of the query

³⁵ <http://www.bing.com/toolbox/bingdeveloper/>

4.3.2.2. Expand

As mentioned in section 4.3.1.2, ARCHING contains a component that generates expanded queries based on (i) an original user query and (ii) textual content that reflects the user's current information interest.

The algorithm used by this component is based on an improved version (Carpineto, et al., 2001) of the original Rocchio (Rocchio, 1971) algorithm for query expansion, which produces a query vector Q_{new} that expands an initial query vector Q_{orig} according to the following formula:

$$Q_{new} = \alpha \cdot Q_{orig} + \frac{\beta}{|R|} \sum_{r \in R} r - \frac{\gamma}{|R'|} \sum_{r' \in R'} r'$$

where Q_{new} is a weighted term vector for the expanded query, Q_{orig} is a weighted term vector for the original unexpanded query, R and R' are respectively the sets of relevant and nonrelevant documents, r and r' are two term weighting vectors extracted from R and R' , respectively. The weights in each vector are computed by a weighting scheme applied to the whole collection.

If the query expansion only relies on positive feedback, the equation reduces to:

$$Q_{new} = \alpha \cdot Q_{orig} + \frac{\beta}{|R|} \sum_{r \in R} r$$

This formula is implemented in an open-source Lucene-based library³⁶ and has been integrated into ARCHING with a custom Java wrapper.

There are several parameters that can be set when executing this component:

- *Query*: the original query string
- *Language*: the natural language of the query
- *Page location*: if the module is to extract text from a particular URL, this parameter specifies this URL location
- *Text*: if text has been extracted *a priori*, this parameter can be used instead of the page location parameter

³⁶ <http://lucene-qe.sourceforge.net/>

- *Index path*: the Lucene index base location of the reference document collection
- *Index name*: the Lucene index name
- *Expansion type*: currently choosing between “Lucene” (stemming) and “Simple” (no stemming). When using open-web search, “Simple” needs to be selected, as the Bing API does not support stemmed queries. If “Lucene” is chosen, query term boost weights are also returned for the stemmed queries.
- *Extra terms*: in some cases it may be desirable to specify particular terms to be added to the query, e.g. if a query relates to a specific product and the search is run across the open web, adding the product name may help disambiguate the query
- *Exclusion terms*: in some cases it may be desirable to exclude certain terms, e.g. spam terms such as “keygen” or “crack”

4.3.2.3. Translate

Similar to open-web search, the ARCHING implementation contains translation functionalities through the Microsoft Translator API³⁷. Custom Java modules have been implemented to fully integrate this RESTful service into the ARCHING architecture.

The parameters when calling this module are:

- *Text*: the text to be translated
- *Source language*: the language of the source text
- *Target language*: the language of the translation result

4.3.2.4. Classify

The classification module in ARCHING uses a Rocchio classifier, which maps documents to precomputed concept centroids³⁸. These centroids are trained from term vector averages of pre-classified documents:

³⁷ <http://www.microsofttranslator.com/dev/>

³⁸ <http://nlp.stanford.edu/IR-book/html/htmledition/rochio-classification-1.html>

$$\vec{\mu}(c) = \frac{1}{|D_c|} \sum_{d \in D_c} \vec{v}(d)$$

where D_c is the set of pre-classified documents whose concept is c . The normalized vector $\vec{v}(d)$ of d is calculated using *tf-idf* (see section 4.3.2.1).

The classification process calculates the cosine similarity (see section 4.3.2.1) between the normalised term vector of the new document and each concept centroid. Concepts with high similarities between their centroid and the new document's term vector can hence be assigned to the document with a certain probability (i.e. the calculated similarity score).

This classifier is implemented in an open-source Lucene-based library³⁹ and has been integrated into ARCHING with a custom Java wrapper. In the current implementation of ARCHING, concepts refer to domain ontology concepts and centroids may be learned from preclassified/annotated closed-corpus content.

4.3.2.5. Additional Design flexibility

As mentioned in section 4.3.1.5, additional flexibility is required for prototypes that need to be able to cope with delays introduced by open-corpus queries. To this end, it is possible to set a search parameter for asynchronous requests, which specifies that the narrative only generates the query and result wire frame, but without running the actual search.

In this scenario, the full result composition is hence generated in two separate steps:

- *Step 1*: AE-based generation of result composition, including query and result wire frames for open-corpus search
- *Step 2*: Asynchronous client-side execution of open-corpus search based on the queries generated in Step 1.

This two-step process ensures a responsive user experience, as compositions can be returned to users without needing to wait for all results to be returned from open-corpus search modules. Figure 4-3 illustrates an example of this process, whereby an initial

³⁹ <http://jtmt.sourceforge.net/>

user query is processed by the Adaptive Engine in order to generate an initial result presentation (Presentation.jsp). This presentation may already contain initial results (either closed-corpora or open-corpora), as well as placeholders for asynchronous result requests.

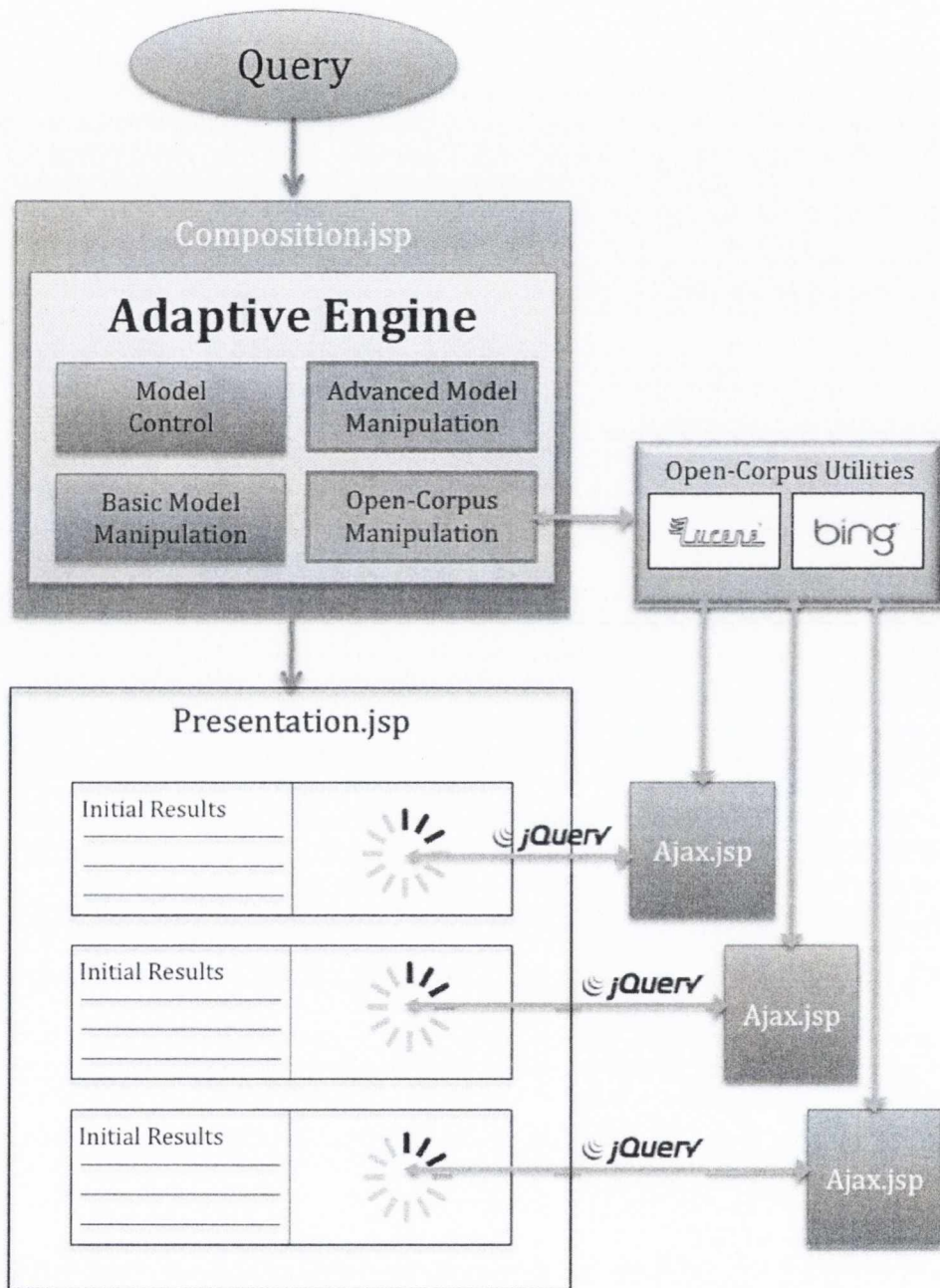


Figure 4-3. Asynchronous Open Corpus Utility calls

This client-side capability has been integrated using the Ajax⁴⁰ functionality of the Javascript library jQuery⁴¹. Ajax requests are issued to custom JSP pages, which have full access to the open-corpus utility functions described above.

4.4. Prototype Implementation

In order to study and evaluate the presented techniques and technologies for adaptive open-corpus retrieval, composition and presentation, an implementation of ARCHING has been applied in a case study of *Personalised Customer Care (PCC)*. This scenario represents a suitable application area, as it provides (i) authentic information needs, (i) heterogeneous data sources as well as (iii) real-life evaluation possibilities.

Section 4.4.1 first introduces the domain of customer care and describes the current challenges in customer and information diversity. Section 4.4.2 describes the content ecosystem surrounding the PCC prototype, including the variety of customer care information sources and their related metadata models. Section 4.4.3 describes the adaptation processes of the prototype implementation, including screenshots of the final composition interface.

4.4.1. Personalised Customer Care

Companies and organisations increasingly face challenges in addressing the various information needs of their customers, particularly given the growing diversity of user experience, context or language preferences. Moreover, in order to successfully establish long-term relationships with their customers, companies increasingly need to be able to provide personalised customer service to attract customer loyalty (Reichheld, 2003). However, many customer support systems have traditionally adopted a simple one-size-fits-all model, leaving users having to either browse through long product manuals or large frequently asked question sections on corporate websites. If a user's search for information using these resources is unsuccessful, a costly customer support call needs to be handled by a customer support agent.

More recently, with the rise of the social web (or Web 2.0), product or service users increasingly engage in third-party community forums in order to solve issues in a

⁴⁰ <http://w3schools.com/ajax/default.asp>

⁴¹ <http://jquery.com/>

community effort. The popularity of this paradigm has motivated many companies to provide their own versions of such forums in order to leverage and exchange knowledge with their user community. However, these community resources are typically held separate from the traditional, corporate content, requiring users to sort through many search result lists of un/semi-structured forum entries.

There remains tremendous potential for combining the complementary benefits of these various information resources, which could prove beneficial for companies as well as their customers. Moreover, by providing better assistance, guidance and navigation for individual users and their needs, companies will be able to improve the increasingly important customer loyalty score by providing a more Personalised Customer Care.

Through the integration of various support data and the provision of adaptive navigations across them, this chapter (as well as the alternative prototypes presented in chapters 5 and 6) illustrates that ARCHING is able to bridge the gap between heterogeneous information sources and diverse user information needs and contexts.

4.4.2. Prototype Ecosystem

The PCC system described in this chapter uses a hybrid of closed-corpus and open-corpus content focussing around the Symantec security product Norton 360⁴².

The closed-corpus data consists of highly structured versions of several product manuals, as well as online help documentations. Using customised scripts, this information (originally formatted in the DocBook Document Type⁴³) has been automatically transformed into reusable semantic knowledge items represented as instances of an OWL ontology (Sah and Wade, 2010) (see Figure 4-4).

⁴² <http://us.norton.com/360>

⁴³ <http://www.docbook.org/specs/cs-docbook-docbook-4.2.html>

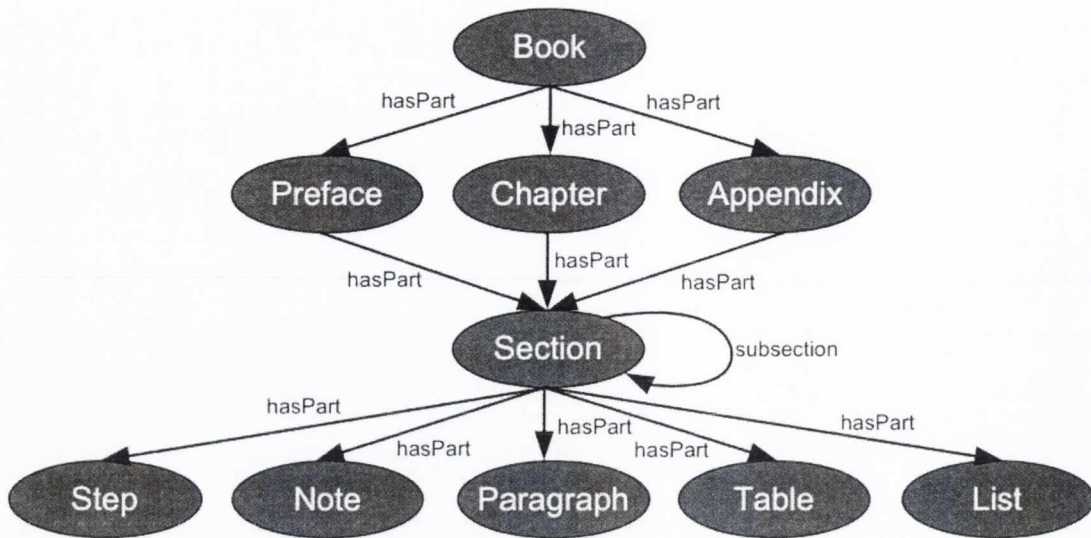


Figure 4-4. Part of the DocBook structure modelled as ontology classes

Additionally, a fuzzy-logic-based metadata generation process (Sah and Wade, 2011) has generated several metadata attributes for personalisation, including the *difficulty* of an item, as well as its *interactivity type* and *interactivity level*. In total, these metadata extraction processes have generated over 600 individual knowledge items and their respective metadata.

While difficulty is typically calculated in combination of a user and a topic, the difficulty level in this case study is based on the complexity (e.g. based on content length) and number of concepts covered in an item (e.g. number of tables, lists). It could therefore be defined as representing an item's level of detail. The difficulty property reuses the Learning Object Metadata⁴⁴ (LOM) vocabulary and ranges from *very easy* to *very difficult*.

The interactivity type and interactivity level refer to the type of information conveyed by a knowledge item, ranging from being more explanatory (expositive material) to more instructional (e.g. guided instructions). To this end, the number of "Procedure" and "Step" elements in the documents are indicators for more *active* documents, whereas elements such as "Table", "List" or "Note" denote more *expositive* documents. As with the difficulty property, the LOM vocabulary has been reused. The possible values for interactivity type are *active*, *expositive* and *mixed*. For interactivity level, values can range from *very low* to *very high*.

Moreover, document index terms (commonly found in such corporate documentation) have been used for the automatic generation of a concept ontology. Manual cleaning

⁴⁴ <http://ltsc.ieee.org/>

has been performed on this ontology in order to remove duplicates and to enhance relationship values between concepts. This semi-automatically generated ontology is used as the high-level domain model of the PCC prototype system (see Figure 4-5).

The index-term/concept-linking ensures that each knowledge item relates to a particular concept in the domain ontology. This enables the semantic retrieval and composition of knowledge items depending on conceptual values. Moreover, this linking can be used in order to train the classifier described in 4.3.2.4, using the knowledge items as the training documents and the index terms as classes.

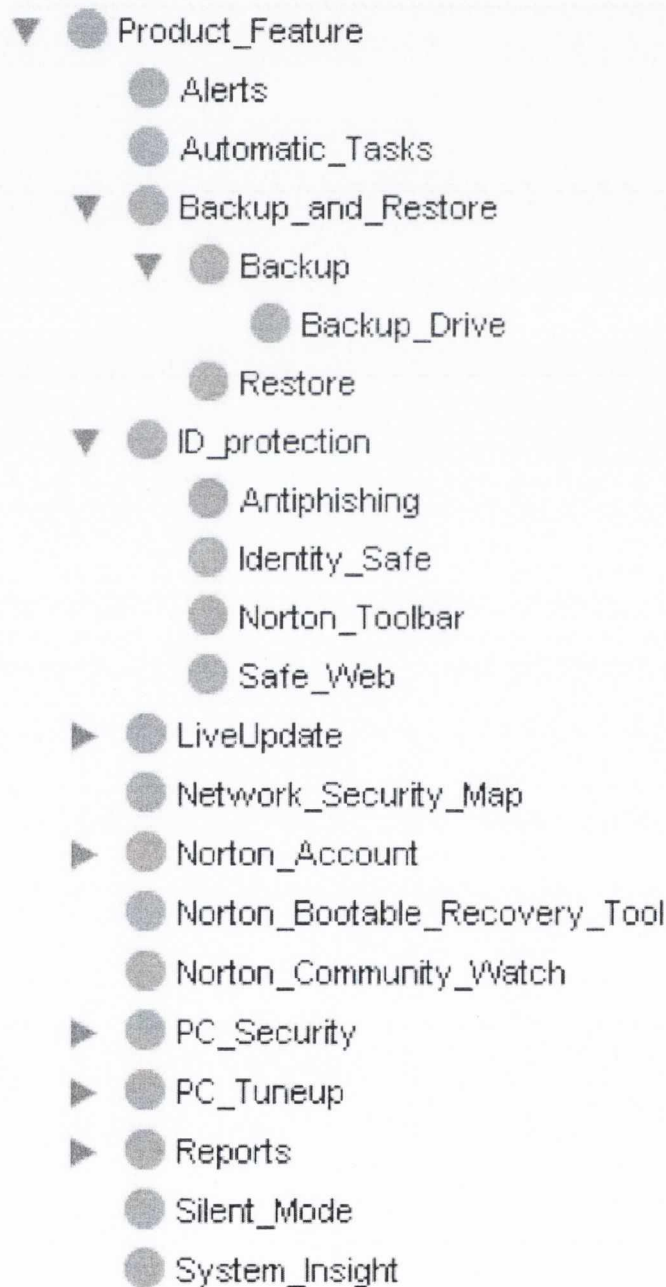


Figure 4-5. Domain ontology of product features

In addition to this closed-corpus (metadata-rich) corporate information, unstructured open-corpus content has been collected using standard harvesting technology. For the purposes of this initial PCC study, user forums (which are maintained independent of the corporate site on a third-party website) have been crawled using the open-corpus harvesting system OCCS (Lawless, et al., 2010). This process has yielded a collection of open-corpus data in excess of 10,000 forum entries. As opposed to the prototype presented in Chapter 3, no additional metadata has been generated for this open-corpus content cache.

It has to be noted that both OCCS and ARCHING are able to handle any content available on the web. However, for the purposes of this initial case study evaluation, only user forums regarding Norton 360 have been crawled and indexed by ARCHING. Also, to maintain a reliable and consistent evaluation corpus, the harvested content has been stored in a local cache. By contrast, chapters 5 and 6 describe alternative ARCHING prototypes, which do not rely on such local caches, as they perform open-corpus searches using the open-web capabilities described in 4.3.2.1.

4.4.3. Adaptation Processes

This section describes the adaptation processes used by the ARCHING-based PCC prototype to generate adaptive compositions across the aforementioned (heterogeneous) data sources, including the structured knowledge items (i.e. product documentation) as well as the unstructured information (i.e. forums).

The PCC system provides two types of compositions, *overview results* and *detailed results*. Both compositions are generated through a process that encompasses a series of adaptation steps. Each of these steps may use a number of models, including user model, domain model, content model, transformation model, as well as closed/open-corpus content indices. These models are consolidated with a narrative in the Adaptive Engine in order to generate result models according to the narrative strategy and the relevant model states and preferences.

Section 4.4.3.1 first describes the various models used in the adaptation processes, followed by a description of the steps involved in the overview result adaptation (section 4.4.3.2) and detailed result adaptation (4.4.3.3).

4.4.3.1. Adaptation Process Models

User Model

In the initial PCC system presented in this chapter, a relatively simple user model consists of the following preferences that are captured during the query elicitation stage (see Figure 4-6 for query elicitation interface):

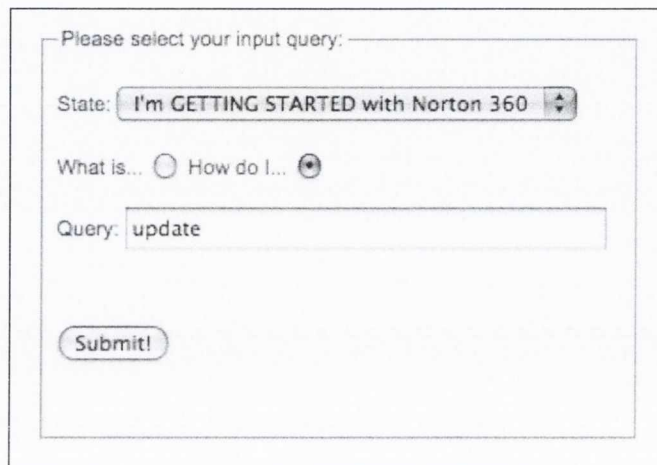


Figure 4-6. Query Elicitation

- First of all, users specify their current state regarding the product, which ranges from “Installing Norton 360”, to “Getting started with Norton 360”, “Reacting to a particular problem.”, etc. This preference represents a user’s current context with respect to the product they are querying about and helps the system select, group and sequence appropriate knowledge items in terms of activity type and difficulty level.
- Secondly, users specify a question type, which consists of either “What is” or “How do I”. This preference defines the query intent, allowing users to specify if they are looking for explanations regarding a particular feature or if they prefer more tutorial-style answers. This preference is used by the system in order to select, group and sequence appropriate knowledge items in terms of activity type and activity level.
- Lastly, users input a free-text keyword query to indicate their information need.

Domain Model

The domain model in the PCC prototype consists of the domain ontology presented in section 4.4.2. This model describes the various product features of the Symantec product Norton 360, including the hierarchical relationships between features.

Content Model

The content model contains the various metadata values held by the structured knowledge items, including the concept related to the content, the difficulty level, as well as the activity type and activity level. It has to be noted that this model is only the content model of the structured data sources (i.e. the product documentation knowledge items). There is no content model marked up for the unstructured open-corpus content (i.e. forums), as such information sources are handled using the open-corpus indexing, retrieval and adaptation techniques presented in 4.3.2.

4.4.3.2. Overview Results - Adaptation Process

Figure 4-7 illustrates the process for generating the overview results, which encompasses a series of adaptation steps in order to generate the first result composition.

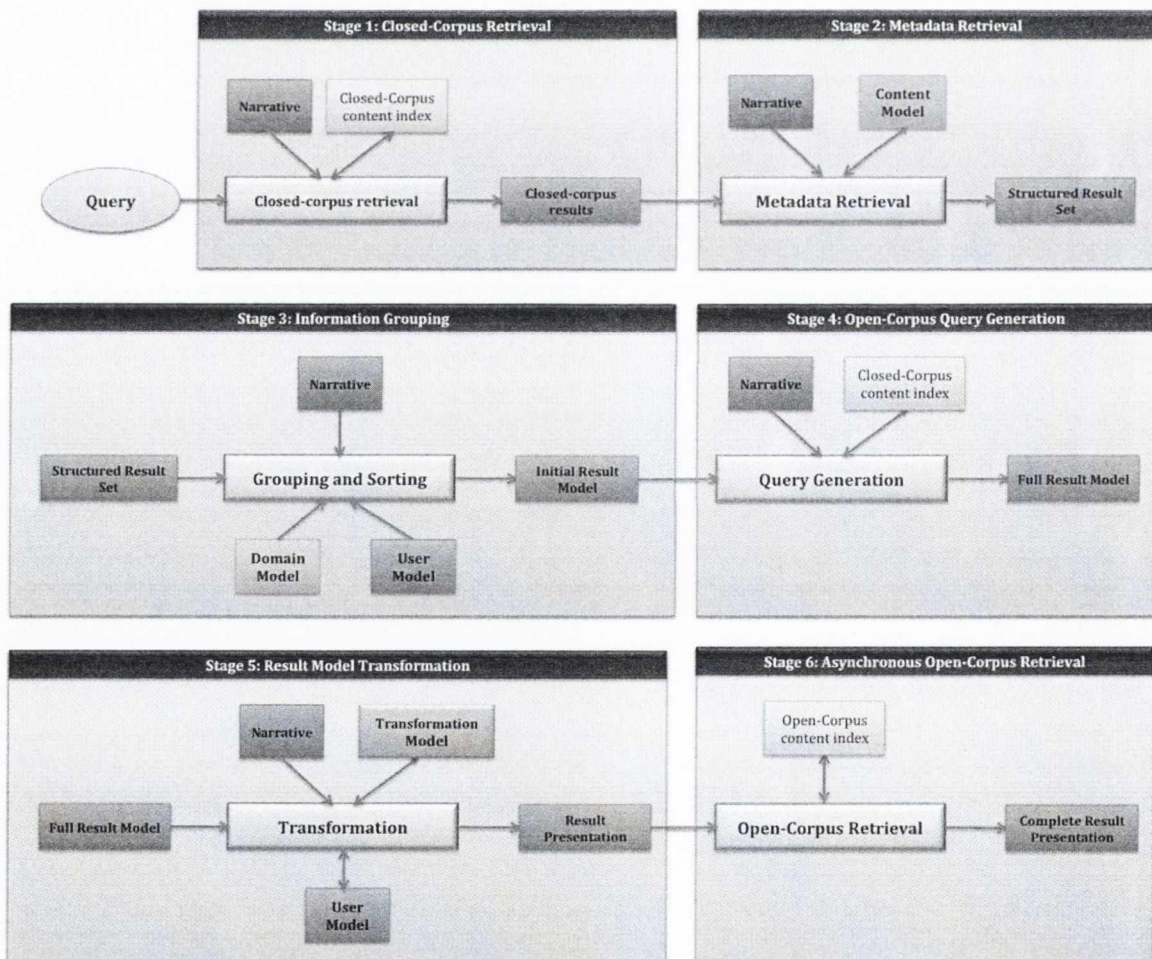


Figure 4-7. Result overview generation process

Stage 1 & 2: Closed-Corpus Retrieval and Metadata Retrieval

In the first stage, the user query is executed on the knowledge item index in order to retrieve an initial result set. The results from this step also contain a relevance score for each retrieved knowledge item based on the statistical similarity described in section 4.3.2.1.

In a second step, metadata values are retrieved in order to determine the conceptual space of the initial results (i.e. their corresponding concepts), as well as their difficulty and activity type/level. At the end of this second step, a structured result set has been composed, which contains knowledge item results, as well as their associated metadata. Figure 4-8 illustrates these initial steps, starting with a user query and resulting in a structured result set model.

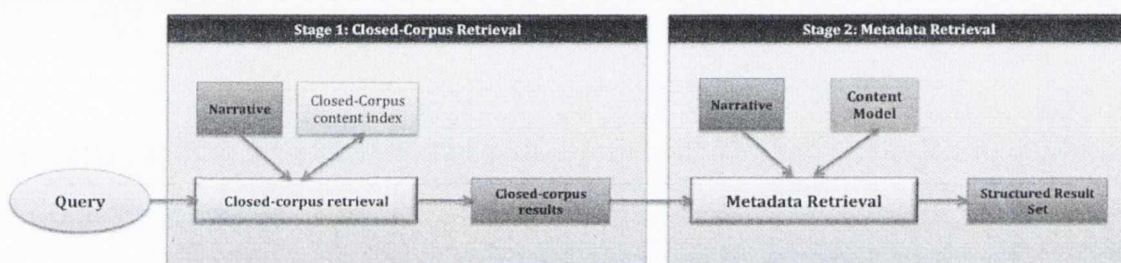


Figure 4-8. Closed-Corpus Retrieval and Metadata Retrieval

Stage 3 & 4: Information Grouping & Open-Corpus Query Generation

Following this initial retrieval, the various knowledge item results are grouped according to their associated concepts. Conceptual scores are calculated based on the aggregate score of the underlying results. Moreover, results within groupings are sorted based on the indicated user state and intent. For example, if a user has indicated “I’m getting started with Norton 360” and prefers “How do I”-type information, the ranking scores are promoted for results that have a difficulty type of “Very Easy” and “Easy”, as well as well as the activity type “Active”.

Also, related concepts are semantically retrieved for each domain concept that is associated with the initial results. This allows later composition stages to (i) provide adaptive navigations across related concepts (stage 5) and (ii) retrieve additionally relevant knowledge items. Figure 4-9 (left) illustrates this third step, which takes as an input the structured result set produced by step 2 and outputs a grouped overview result set (called *Initial Result Model*).

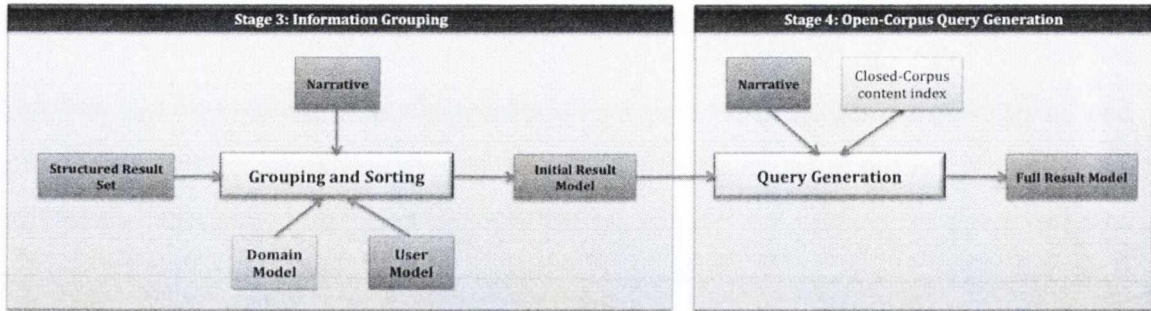


Figure 4-9. Information Grouping & Open-Corpus Query Generation

In the next step (stage 4), for each concept contained in this initial result model, the complete textual content of the top ranking knowledge item is used together with the original query as the input for the statistical query expansion component. The resulting adapted queries are then added to the initial result model, in order to create the *Full Result Model* (which is now ready for display transformation). This model hence contains i) the initial result model generated in stage 3, and ii) an associated query for each of the concepts in this model.

Stage 5 & 6: Result Model Transformation & Asynchronous Open-Corpus Retrieval

The final result grouping (stored at the end of stage 4) is transformed to XHTML using XSLT in order to form the first interlinked overview result presentation, which is then displayed to the user (see Figure 4-10).

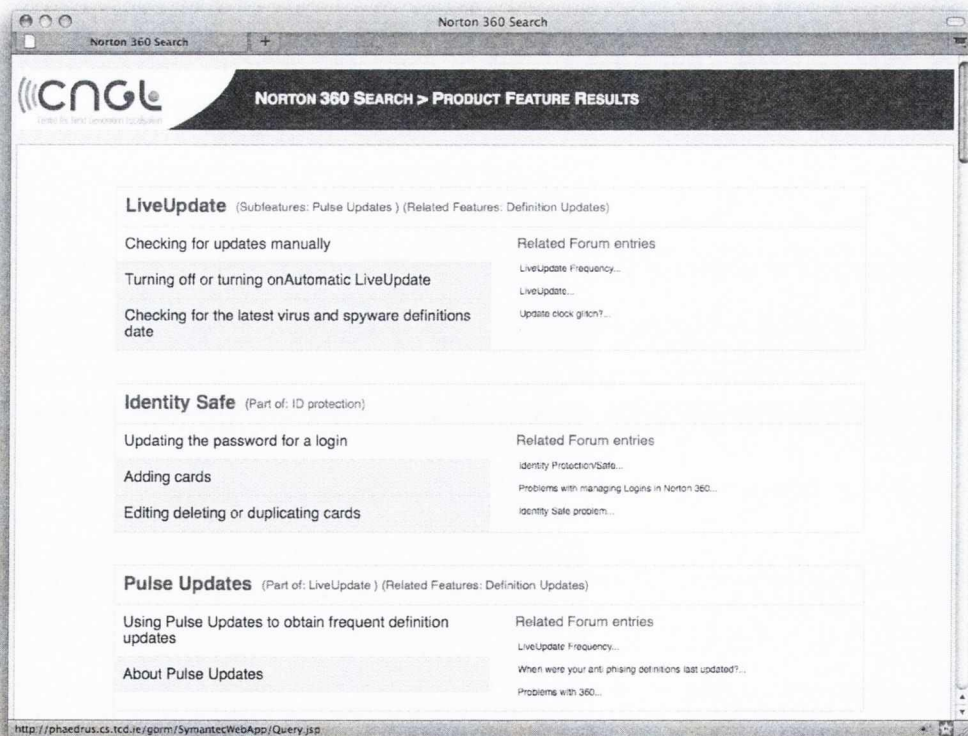


Figure 4-10. Result Overview Screen

This overview screen provides users with a structured presentation of information, with concepts being used as group headers. This enables users to disambiguate the various results, as in the given example, there are many interpretations to the user query “update” (e.g. updating virus definitions through “LiveUpdate”, updating “Identity Safe” passwords, receiving “Pulse Updates”).

Moreover, this screen provides users with an interlinked conceptual space, allowing them to identify relationships between the various concepts. For example, Figure 4-11 shows the top-ranking concept following the query “online backup”. In this example, the “Backup” feature is part of a feature called “Backup and Restore” and has a subfeature called “Backup Drive”. By clicking on the respective links, the screen scrolls automatically to the grouping of these features (if there are results for these concepts).

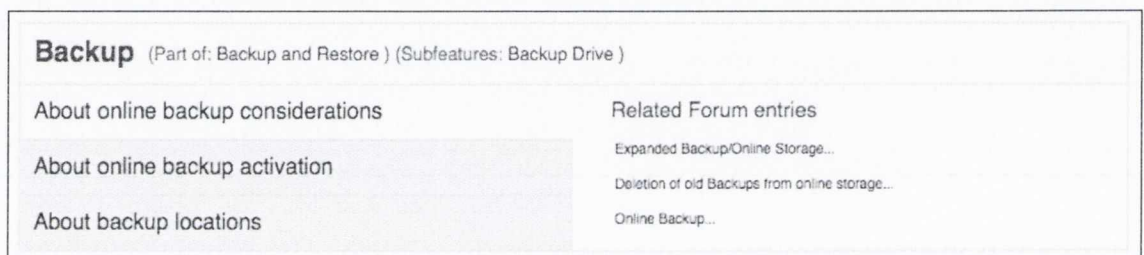


Figure 4-11. Overview Results for “Backup”, including related features (top)

As can be seen in both Figure 4-10 and Figure 4-11, the result groupings contain documentation results (left), as well as related forum entries (right). This additional (open-corpus) forum information is retrieved asynchronously (see Figure 4-12) using the associated query generated during stage 4 (i.e. the query that is associated with the top ranking knowledge item for this concept). The expanded queries are executed on the indexed forum content, producing ranked result sets that are composed into the final result presentation.

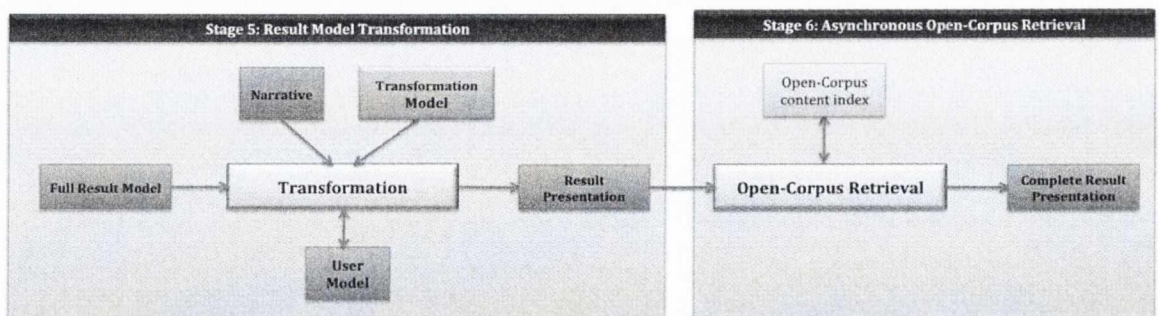


Figure 4-12. Result Model Transformation & Asynchronous Open-Corpus Retrieval

This integration enables a meaningful combination of the unstructured content into the highly structured, grouped and ranked set of documentation results. As forum entries are retrieved in the same manner for each of the conceptual groupings, the unstructured content is effectively integrated into a structured hypertext.

Moreover, when a user hovers over the second or third documentation result, a new query is generated on-the-fly (using the textual content of this knowledge item) and executed on the open-corpus index. This enables users to get an initial overview across the forum content using the topics provided in the documentation.

4.4.3.3. Detailed Results - Adaptation Process

After a user selects one of the results (either a knowledge item from the documentation or a forum entry), a more detailed presentation is generated for the chosen result's concept (and its related concepts) using a similar process (see Figure 4-13). This detailed result composition consists of all the results that were retrieved for the given concept.

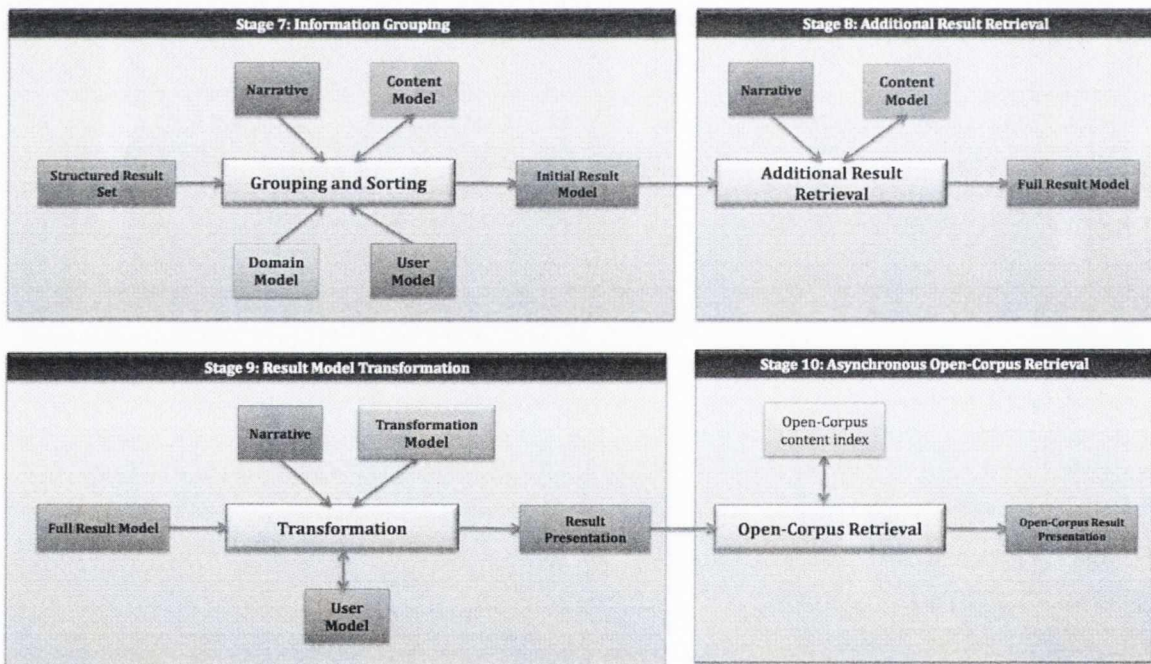


Figure 4-13. Detailed Results - Adaptation Process

Stage 7 & 8: Information Grouping & Additional Result Retrieval

In stage 7, the structured result set produced at the end of stage 2 is grouped to a more detailed level, namely according to the activity type and difficulty level. Also, if

knowledge items in a grouping are related in the content model (e.g. sections/subsections), more fine-grained sub-groupings are composed. Sub-groupings and results within sub-groupings are ranked according to their original keyword score.

The relevance of groupings depends on the user state/intent and the system assists users towards the most suitable content using these values. For example, if a user has indicated to be “Getting Started with Norton 360” and has chosen the question intent of “What is”, the “expositive” and “easy/very easy” content will be marked as the currently most suitable grouping. While this step only identifies and marks the most suitable grouping (by adding an attribute in the result model), the actual realisation (visualisation) of this highlighting occurs later during the result transformation stage.

In step 8, additional knowledge items are retrieved (using the content model) for the concepts that are related to the chosen concept (see Figure 4-14).

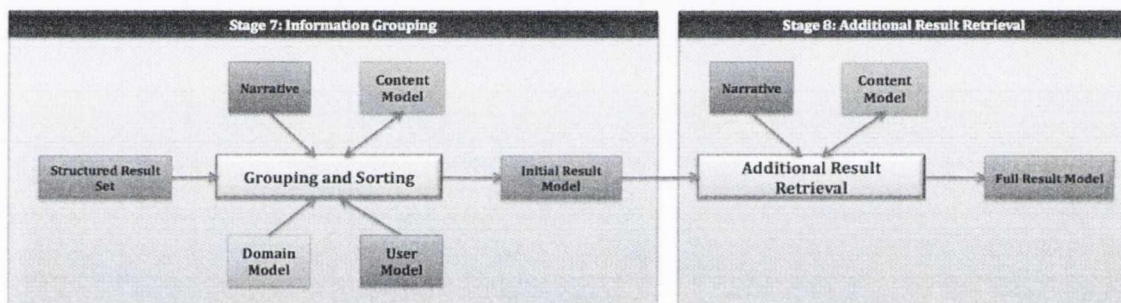


Figure 4-14. Information Grouping & Additional Result Retrieval

For example, if the user has chosen to view a result related to the “Backup” feature, introductory explanations for “Backup and Restore” and “Backup Drive” are retrieved and composed into the full result composition.

Stage 9 & 10: Result Model Transformation & Asynchronous Open-Corpus Retrieval

The result grouping is again transformed into an Adaptive Hypertext using a transformation model. Figure 4-15 shows an example of a structured navigation following a user’s selection of the knowledge item “About online backup considerations” from the original overview results shown in Figure 4-11. In this particular example, the user has previously indicated that she is “Getting Started with Norton 360” and has given a question intent of “What is”. On the left-hand side, a tree-based navigation shows the composition of the various results (see Figure 4-16 for a more detailed view).

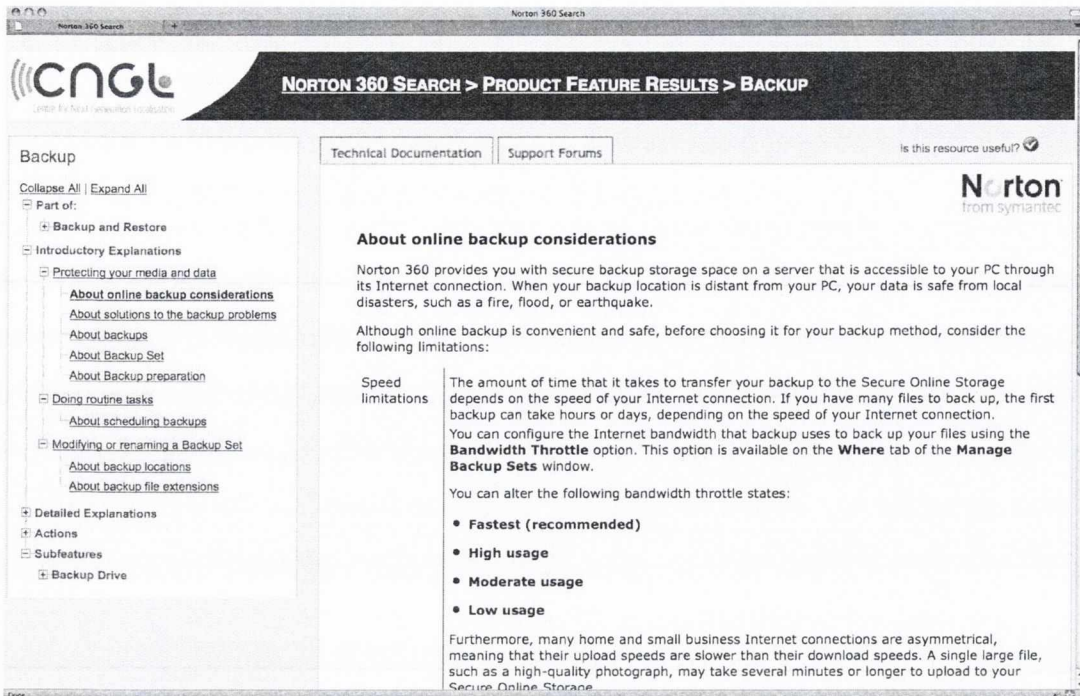


Figure 4-15. Detailed Results for the “Backup” feature, currently displaying knowledge item from documentation

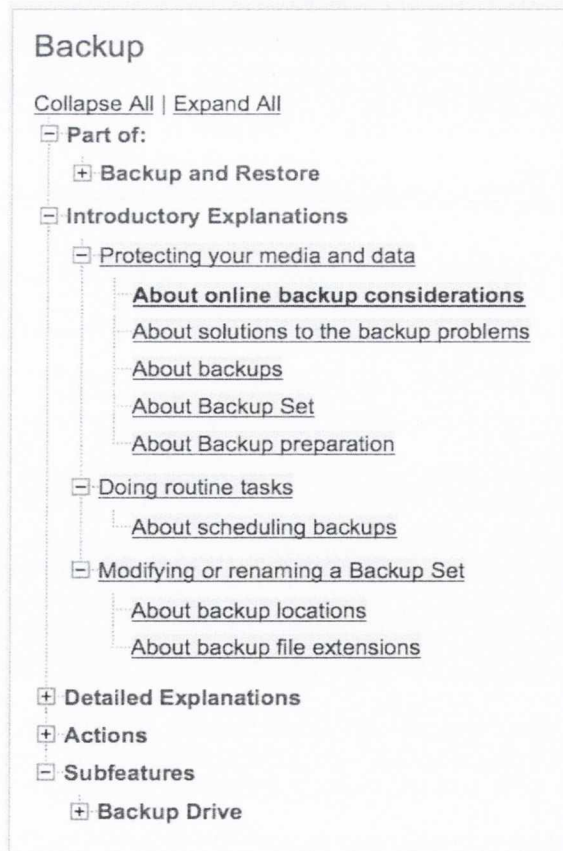


Figure 4-16. Structured Navigation

As can be seen from these figures, the “Introductory Explanations” grouping (previously marked as the currently most suitable context) is already opened up for the user. Under this grouping, most information is of an introductory nature, allowing beginners to learn about basic features before learning about how to configure more complex functionalities. The composed tree also contains the various other results, ranging from more detailed explanations to “Actions” (i.e. tutorial-style information), as well as to related results (i.e. “part of”, “subfeatures”).

Similar to the result overview screen, the unstructured forum content is composed into this structure using query expansion techniques. When a user clicks on the “Support Forums” tab, an expanded query is generated on-the-fly using the original user query and the text from the currently selected knowledge item (see Figure 4-17).

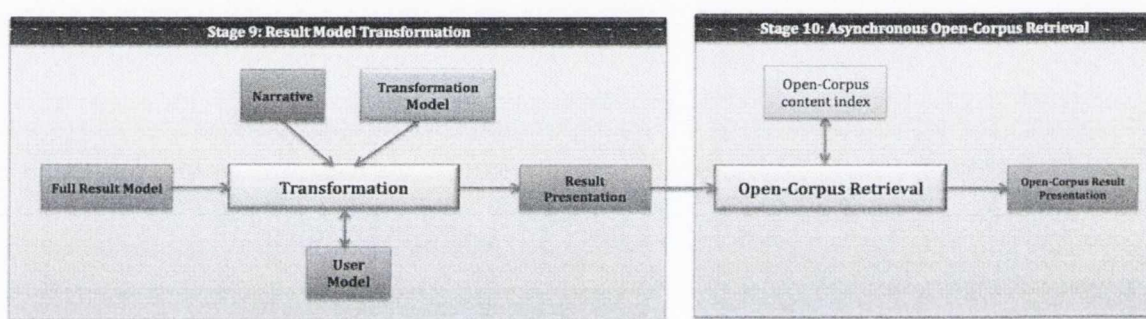


Figure 4-17. Result Model Transformation & Asynchronous Open-Corpus Retrieval

This asynchronous call results in a ranked list of forum entries that are related to the current context of the user (i.e. currently selected knowledge item) (see Figure 4-18).

While having selected the “Support Forums” tab, a user can also use the structured navigation on the left in order to refine the original query towards other topics. For example, if the user selects the knowledge item “About scheduling backups”, a new query is generated on-the-fly, resulting in a ranked list that is targeted towards this topic.

This combination highlights again the open-corpus integration capabilities of ARCHING, as previously separate information sources are combined in an adaptive composition, allowing users to navigate seamlessly across heterogeneous data. Moreover, the previously unstructured forum entries can now be navigated using a fully adapted navigation by making full usage of the corporate knowledge provided in the knowledge items as well as the user query, state and intent.

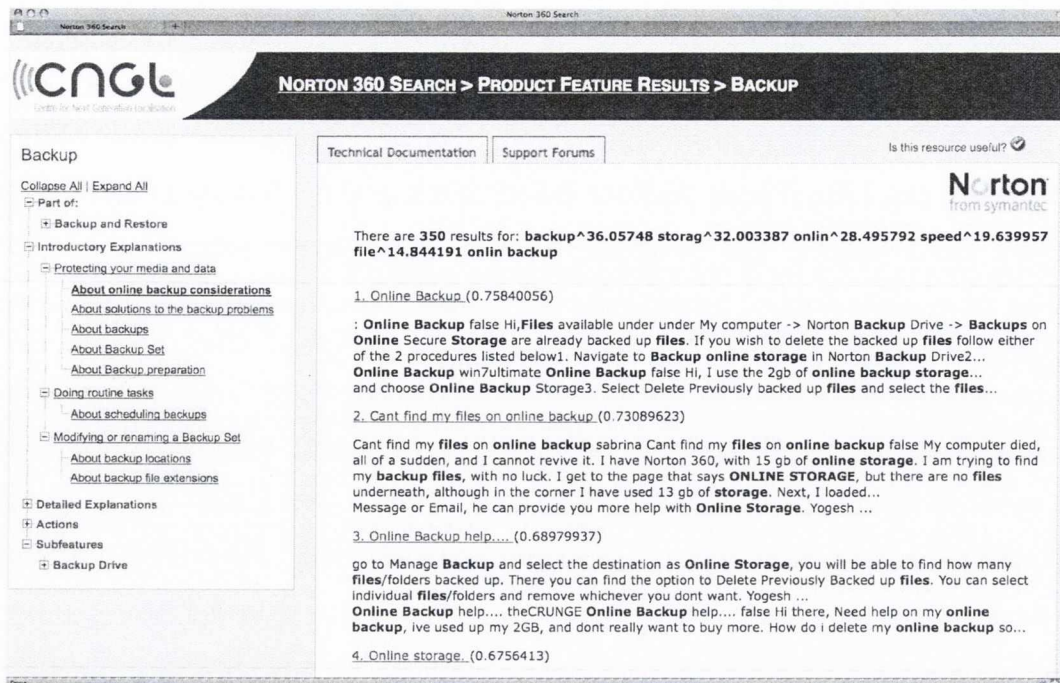


Figure 4-18. Detailed Results for the “Backup” feature, currently displaying support forums

4.5. Evaluation

In order to evaluate the adaptive open-corpus composition approach provided by ARCHING, a real-life user-study was performed using authentic information needs in the context of customer support.

In particular, the presented Personalised Customer Care prototype was evaluated in terms of (i) the ability to support users in real-life customer support tasks (user efficiency and effectiveness) and (ii) the usability from the users’ perspective (i.e. user satisfaction).

4.5.1. Hypotheses/Sub hypotheses

Task Assistance

Compared to the educational case study evaluation presented in section 3.5, the benefit to the user in a customer support scenario lies in a system’s ability to assist a user’s search for information effectively and efficiently. In particular, it is desirable that a system requires users to invest the least amount of effort in order to find relevant information as quickly as possible.

The hypotheses regarding the user efficiency and effectiveness are therefore as follows.

- H1: An Adaptive Information Composition system better assists a user's search for information than a non-adaptive search system.

- H1.1: The adaptive system allows users to be more efficient in terms of user effort for task completion.

The metrics used to test this hypothesis are completion time and number of queries issued.

- H1.2: The adaptive system allows users to browse and view more relevant information.

The metric used to test this hypothesis is the users' overall page view count.

- H1.3: The adaptive system allows users to be more effective for task completion than a non-adaptive system.

The metric used to test this hypothesis is the users' measured and perceived task accuracy.

The corresponding null hypothesis for H1.1-H1.3 is that there are no differences between the adaptive system and the non-adaptive system.

User Satisfaction

In terms of user satisfaction, the benefit to users lies in the perceived usability of the various functionalities provided by the Adaptive Composition system. In particular, the assumption is that users perceive the adaptive system to be more helpful for completing the given tasks and that the various functionalities are recognised and valued. The hypotheses regarding the user satisfaction are therefore as follows.

- H2: Users are more satisfied with an Adaptive Information Composition system compared to a non-adaptive search system

- H2.1: The adaptive system outperforms the non-adaptive system in terms of usability.

Usability questionnaire scores are used to test this hypothesis.

- H2.2: Users recognise and value the composition, adaptation and personalisation aspects of the adaptive system

Usability questionnaire scores are used to test this hypothesis.

- H2.3: Using adaptive compositions motivates users to navigate across more content.

Usability questionnaire scores, as well as users' page view counts are used to test this hypothesis.

Again, the corresponding null hypothesis for H2.1-H2.3 is that there are no differences between the adaptive system and the non-adaptive system.

4.5.2. Comparison to Baseline

In order to test the above hypotheses, it was again necessary to provide a baseline search system as a comparison. In the presented customer support scenario, typical corporate search systems are built using information retrieval engines such as Lucene in order to provide users with keyword-based query interfaces. These systems typically present result lists that are ranked according to keyword similarities between the documents and the user query. For the presented user study, it would have been unfair to compare the adaptive system to standard web search engines (e.g. Google), as such standard search engines could only identify the product manuals in their entirety and would not be able to identify individual parts. Consequently, the adaptive system would have gained an unfair advantage due to its improved corporate content base (which has been previously identified as a general problem of web search engines (White, 2007)).

To provide a better and more competitive baseline, the adaptive system was compared to a purpose-built search system, which (i) used the exact same underlying indexing and retrieval models and (ii) operated across the same content base (see Figure 4-19).

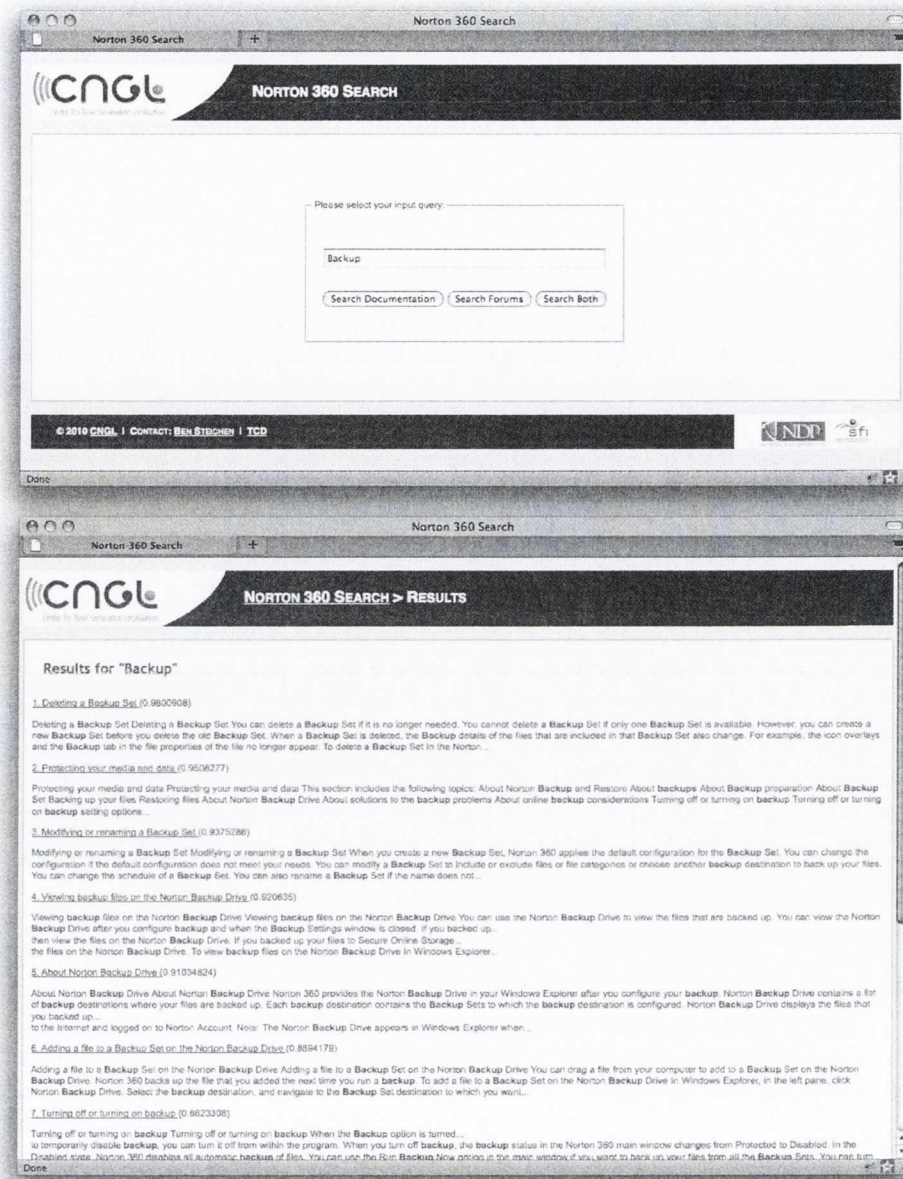


Figure 4-19. Non-adaptive search system

This (non-adaptive) system allowed users to issue a free-text query, as well as to choose between the different content sources (including the option to issue the query across all underlying documents). The results in this search system were ranked according to keyword relevance, hence simulating the conditions in a real corporate IR system typically used by users for the given information seeking tasks.

4.5.3. Experimental Setup

The user-study consisted of a task-based evaluation using real-life information needs regarding the Symantec product Norton 360. An analysis of Symantec customer care

data⁴⁵ provided a set of frequently occurring customer support interaction topics, which were used as the basis for the 4 experimental user tasks. Each task consisted of a set of 3 to 6 questions regarding product features, common problems and possible solutions. Questions ranged from single-answer questions such as “What feature is responsible for updating definitions?” to multiple choice questions such as “Which of the following features are provided by Norton Safe Web”, as well as more open-ended questions such as “Summarise general information on the main functions of the backup feature.” or “You get the following error...Can you find possible causes and/or solutions to the error message (up to two)?”. (The exact task questions, as well as the questionnaires can be found in APPENDIX B.)

The process for each user started by receiving an e-mail about the purpose and length of the experiment, as well as the experiment URL and personal credentials. After logging in to the experiment system, users were asked to fill out a consent form, followed by a pre-questionnaire in order to determine their background regarding the Symantec product Norton 360, as well as their experience with search systems and adaptivity.

Users then received instructions on how to use the first search system (displayed as search system A), including a short video tutorial and the chance to test the system using a test task. After users felt confident with the functionalities of system A, they would then proceed to the first task screen (displayed as task A). This screen consisted of a set of questions, which users were asked to answer using the provided text boxes. In order to solve the task, users were given a link to system A to allow searching across the various content sources. Following the completion of this first task, users were asked to fill out a SUS questionnaire (Brooke, 1996), as well as an application-specific usability questionnaire using a 4-point Likert scale (1=strongly disagree, 2=disagree, 3=agree, 4=strongly agree). This 4-point scale was used deliberately to enforce a user preference.

Following this, users received instructions on how to use the second system (displayed as search system B), including a short video tutorial and the chance to test the system using a test task. They then proceeded to perform a second task (displayed as task B) and filled out the usability questionnaires for system B.

⁴⁵ The data provided by Symantec consisted of training material for customer care agents (including revision questions), as well as logs of real-life customer queries.

Lastly, a final questionnaire directly asked for comparative opinions regarding search systems A and B. After completing the experimental process (see Figure 4-20), users were entered into a random draw for the chance to win an electronic device.

In order to balance any effects of order bias, each user was assigned with two out of the four possible tasks using Latin square design. Also, system order was randomised to ensure that overall the non-adaptive and the adaptive system appeared equally as often as the first system (system A) or second system (system B).

The experimental process was entirely online and users were asked to perform the experiment in a single session without interruption. User actions were tracked throughout search sessions in order to be able to analyse users' system interaction behaviour. Also, task completion times were tracked between the first display of a task's questions until a user's submission of the task answers.

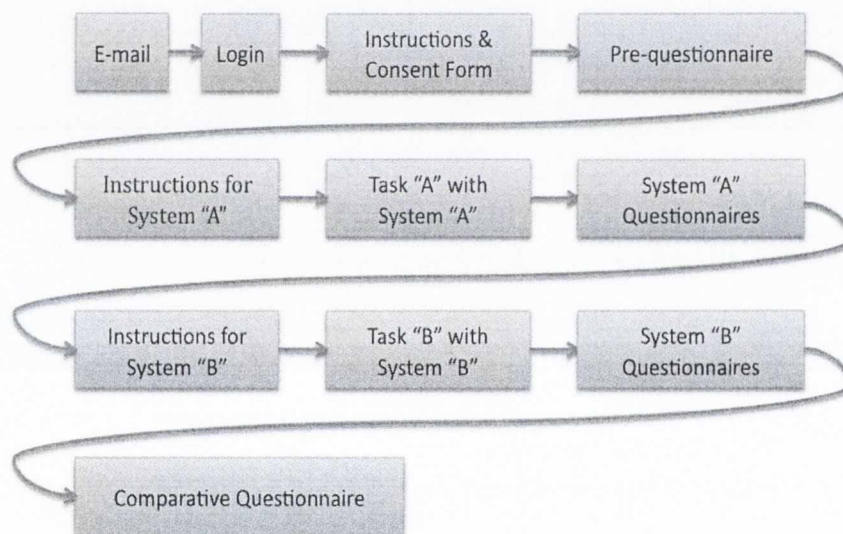


Figure 4-20. Experimental Process

4.5.4. Results

A total of 36 users were recruited from the School of Computer Science and Statistics in Trinity College Dublin and the School of Computer and Information Science at the University of South Australia. The pre-questionnaire revealed that there was little difference between users regarding their Norton 360 experience, stating that they had little knowledge regarding the various product features. In terms of search experience, all users stated that they often use web search engines to search for information about software features in general (medium-high frequency) and that they often consult user

forums rather than product manuals to find problem solutions. When asked about the use of adaptive systems in the past, most users indicated that they had little or no experience.

Task Assistance Results (H1)

As stated in hypothesis H1, the goal of the adaptive system is to better assist users in information seeking tasks. First of all, the results from the task completion times reveal that the adaptive system outperformed the non-adaptive system with an average of 21:54 (mm:ss) versus 25:32. Moreover, paired t-tests confirm that the results are indeed significant ($p=0.031$). Figure 4-21 shows the average times across tasks, showing that users were consistently faster using the adaptive system in tasks 1, 3 and 4.

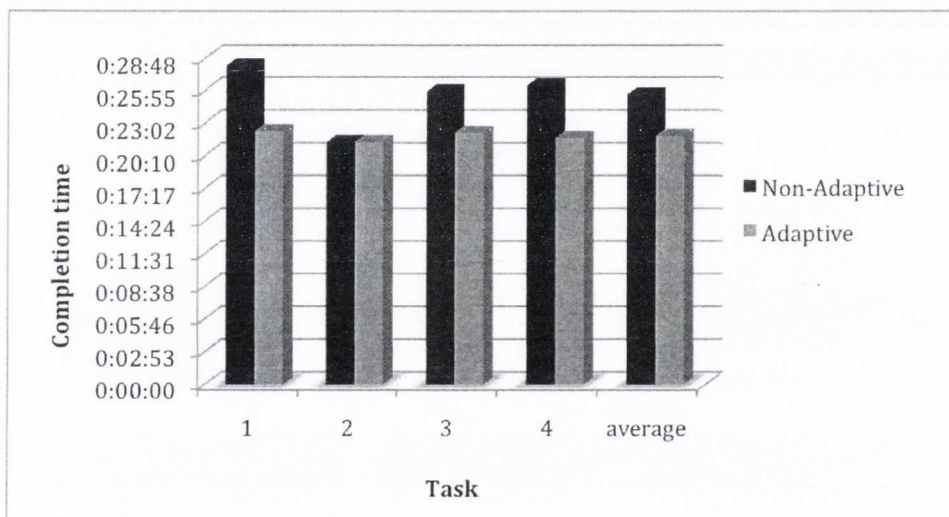


Figure 4-21. Task Completion times

It is interesting to note here that these results are different from the findings in Chapter 3, where users of the AH system spent significantly more time on their (educational) tasks. This difference is further dealt with in the discussion in section 4.5.5.

Similarly, users formulated fewer queries in order to find their information (see Figure 4-22). In the non-adaptive system users required on average 12.09 queries to complete the tasks, whereas the adaptive system recorded an average of 7.25 queries ($p<0.001$).

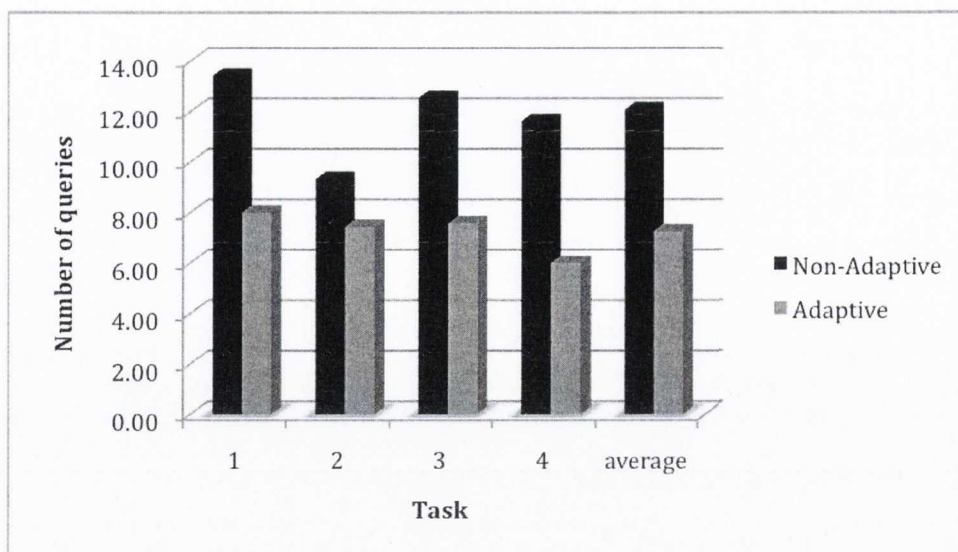


Figure 4-22. Number of queries

Both of these findings clearly point towards the validation of hypothesis H1.1, as shorter completion times and fewer queries reduce the required user effort. The findings are also backed up by open comments from the usability questionnaires, where users reported that in the non-adaptive system they needed to reformulate their query more often in order to reach the desired information. As reformulating queries entails constant result ranking changes, the adaptive system relieved users of having to constantly reorientate. A related questionnaire question asked users if they agreed with the statement “I had to search a lot before I found interesting content”. After using the non-adaptive system, 47.22% agreed, whereas for the adaptive system only 27.03% agreed. Asked directly if they “had to search more” than in the other system, the majority of users agreed for the non-adaptive system (average of 2.69), whereas users mostly disagreed for the adaptive system (average of 2.29) ($p=0.028$).

Another important aspect of task assistance is the ability to allow users to browse and view more information from the various data sources (H1.2). By comparing the click histories across systems and tasks, the evidence shows that users browsed and viewed more information in the adaptive system with an average of 22.31 views compared to the non-adaptive system with 16.51 views ($p=0.002$) (see Figure 4-23). Moreover, in the open comment section of the questionnaires, users confirmed these results with comments such as (in the adaptive system) “I spent less time querying and more time browsing” or “I was less exposed to irrelevant content”.

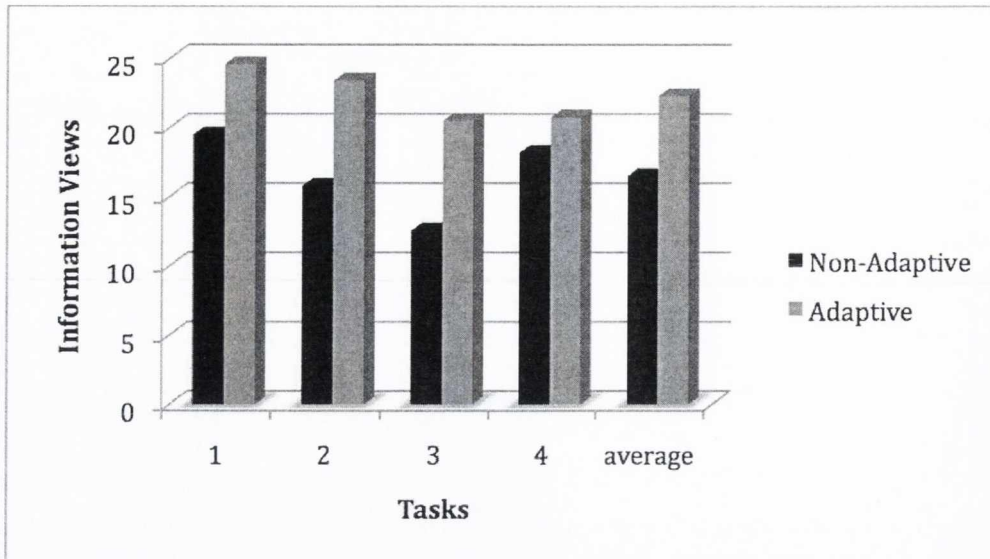


Figure 4-23. Information viewed

When asked to compare the two systems directly if they found that the “search system returned relevant content more prominently”, the majority of users agreed for the adaptive system (average of 3.29) and disagreed for the non-adaptive system (2.05) ($p=0.002$). These results, together with the number of page views, confirm that the adaptive composition and presentation provided by the adaptive system allowed users to navigate more efficiently across the various information sources. Moreover, this finding confirms that the increased page views were not due to the adaptive system presenting more irrelevant results.

In terms of task effectiveness (H1.3), users of both systems scored similar, high results and rarely failed to find accurate information, i.e. for both systems less than 10% of tasks were answered inaccurately. This suggests that most users searched as long as they needed to in order to find the correct answers. However, when asked in the questionnaires if they “did well on the different task questions”, a higher percentage of users agreed or strongly agreed (81.08%) for the adaptive system than for the non-adaptive system (63.89%). Asked directly if they “did better on the different task questions”, the majority of users agreed (average of 3.06) for the adaptive system and mostly disagreed for the non-adaptive system (2.17) ($p=0.010$). These results indicate that the adaptive system made users more confident during the tasks, which is backed up by the majority of adaptive system users agreeing that the “result structure was helpful in solving the tasks” (average of 3.12 versus 2.76) ($p=0.032$). Again, when asked to compare both systems, users agreed that the adaptive system was more helpful (average of 3.12) and disagreed that the non-adaptive system was more helpful (2.05) ($p=0.013$). Although the evidence for the adaptive system did not show higher task

scores, the overall results for H1.3 hence reveal that the adaptive system was at least as effective as the non-adaptive system, while allowing users to be more confident during their tasks.

Overall, the results regarding task assistance clearly indicate the benefits of adaptive compositions and navigations in terms of task efficiency, effectiveness and user confidence. The interaction tracking has provided evidence that users are faster to complete their tasks, require fewer query reformulations and are able to view more information overall. These results have also been backed up by various related questionnaire questions, indicating that users had to search less in order to find relevant information, found the adaptive system to return relevant content more prominently, felt more confident and found the adaptive system to be more helpful for solving the tasks.

User Satisfaction Results (H2)

In addition to the task assistance, the user study aimed at identifying users' appreciation and satisfaction regarding the various functionalities provided by the adaptive system.

First of all, in order to determine the overall usability (H2.1), standard usability scale (SUS) scores were calculated for both the adaptive and non-adaptive systems. In this independent usability score the adaptive system scored an average of 77.35, whereas the non-adaptive system scored an average of 70.37 ($p=0.012$). This is a very encouraging result for such a novel system, especially considering that most users had not used adaptive systems in the past.

This finding is further confirmed by users' answers to the questionnaire question "Overall, I am satisfied with the system performance, assistance and guidance.", with users giving an average score of 3.26 for the adaptive system compared to 2.46 for the non-adaptive system ($p<0.001$). Users gave even stronger evidence for the overall usability of the adaptive system when asked to compare the systems directly. The majority of users agreed/strongly agreed (35.29%/41.18%) that they were more satisfied with the adaptive system (average of 3.25) and disagreed/strongly disagreed (42.11%/31.58%) that they were more satisfied with the non-adaptive system (average of 1.94) ($p<0.001$) (see Figure 4-24).

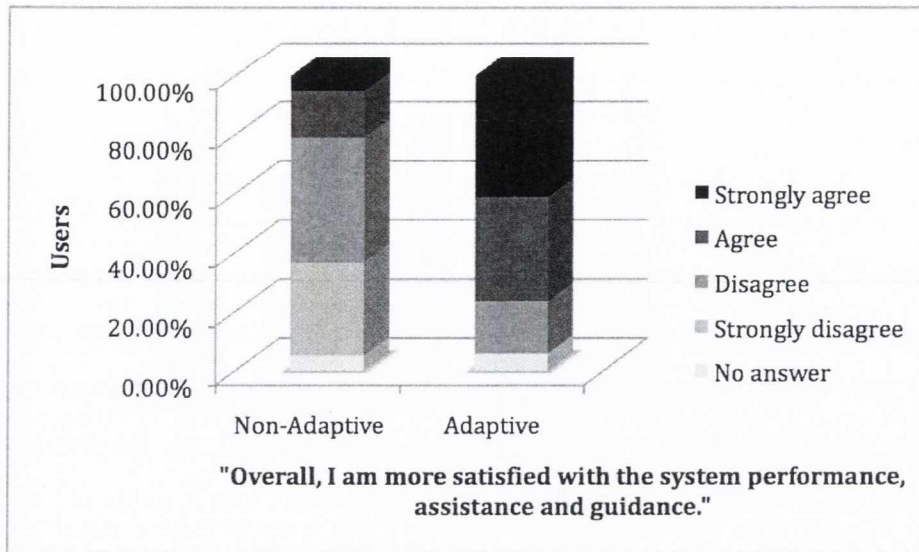


Figure 4-24. User satisfaction

In addition to these more general usability questions, users were asked about the various adaptation, composition and personalisation aspects of the systems (H2.2). These questions were aimed at evaluating more specifically if users recognise and value the various adaptive functionalities. Figure 4-25 provides an overview of the comparative answers given by users. The solid black bars indicate answers for the non-adaptive system, whereas the grey bars indicate answers for the adaptive system.

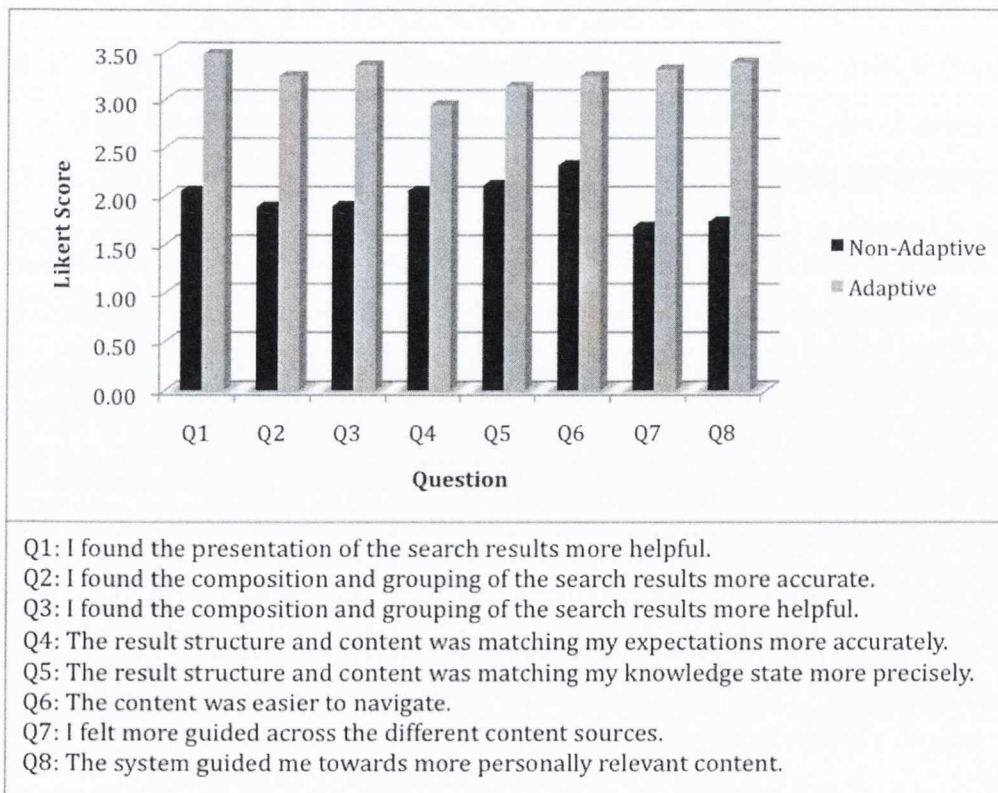


Figure 4-25. Comparative user perception of result presentation, composition and personalisation

The results clearly indicate user preferences towards the presentation, composition and grouping of the adaptive system (Q1-Q6) ($p \leq 0.001$). This again is a very encouraging result given the assumed familiarity of users regarding typical search systems represented by the non-adaptive system. Questions 7 and 8 highlight in particular the strengths of the approach taken by the adaptive system, as users recognise and value the additional guidance provided across the various content sources.

In order to gain more insight into which features were particularly useful, users were asked “what features/characteristics did you like most about the system”. For the non-adaptive system, the dominant responses were its speed (mentioned by 9 users), simplicity (6 users) and the ability to directly query a particular information source (2 users). For the adaptive system, users mentioned that they particularly liked the integration of content sources (20 users), the grouping of results (11 users) and the overall navigation (4 users). Also, some users particularly expressed their liking of the “How/What” personalisation (5 users) as well as the user state personalisation (3 users). In turn, when asked about “what features/characteristics did you like least about the system”, for the non-adaptive system, users mentioned the missing integration of content sources (5 users) and the lack of overall result structure (5 users). For the adaptive system, some users mentioned that they did not like having to manually enter question type and user state (3 users). These answers are certainly encouraging in terms of general user acceptance of adaptive information compositions, while leaving a certain amount of work to be done in terms of implicitly capturing the various user characteristics.

Finally, an important aspect of adaptation and personalisation is the effect of motivating users to engage more with the system (H2.3). The findings regarding user efficiency have provided clear evidence that users are motivated to view more pages, which is backed up by the questionnaire questions regarding motivation, engagement and fun (see Figure 4-26). Most users agreed with the statement “I found the interaction more motivating/engaging/fun” for the adaptive system, whereas the majority of users disagreed for the non-adaptive system ($p \leq 0.001$). These results are again very encouraging as it confirms that one of the main benefits of adaptivity lies in the ability to motivate and engage users to interact more with information systems.

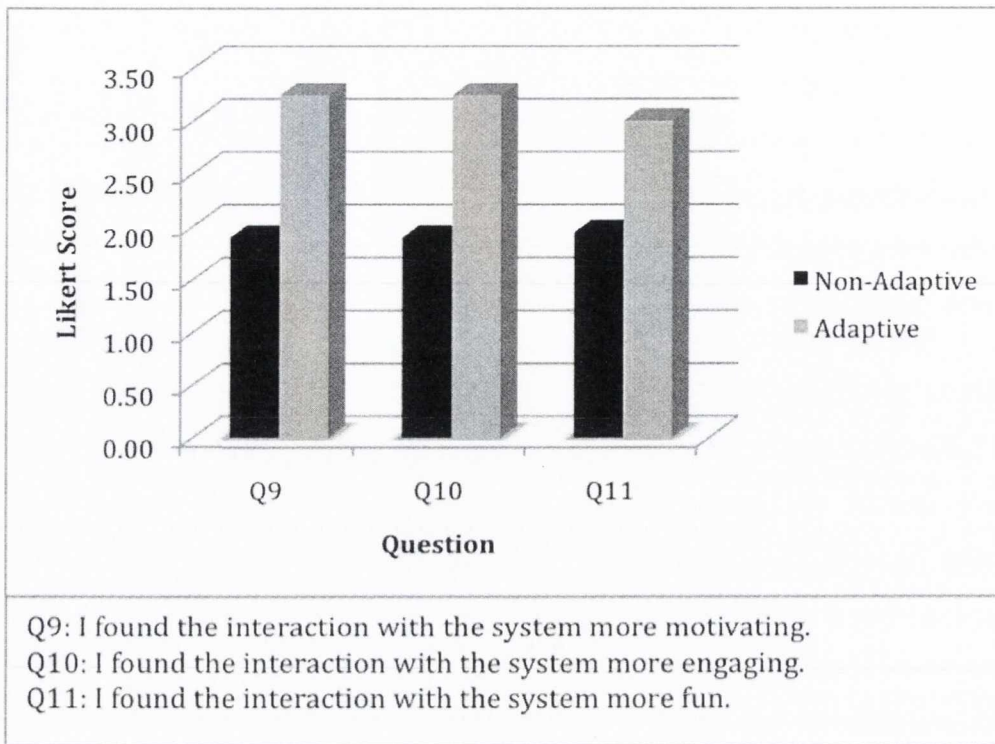


Figure 4-26. User motivation, engagement and fun

4.5.5. Discussion

The results of the user study have revealed a number of encouraging results for providing adaptive compositions across heterogeneous data sources. First of all, it has been shown with evidence that users are effective and efficient at performing information gathering and problem solving tasks. This finding also confirms that the compositional approach not only assists users in educational learning scenarios (such as the case study presented in chapter 3), but also in domains where the user focus lies on *quickly* searching for information and assimilating relevant knowledge. Whereas users spent more time (than non-adaptive users) using the (educational) AH system in chapter 3, the adaptive system in this chapter was able to provide relevant information without requiring users to spend more time on their tasks (which is desirable in the given customer care scenario). This effect has been provided by adding adaptive information retrieval capabilities, allowing users to specify precise information needs in the form of free-text keyword queries. Moreover, by maintaining the multimodel adaptation functionalities, the PCC prototype has been able to assist and guide users towards personally relevant content through adaptive navigation and presentation techniques. This has enabled users to browse efficiently through relevant information, resulting in increased information views despite shorter task times.

In addition to user effectiveness and efficiency, a range of user satisfaction metrics have shown that users recognise and value the various adaptive features provided by ARCHING. General user satisfaction (through standard usability questionnaires) has been increased for the adaptive system compared to the more established information retrieval (ranked-list) delivery paradigm. Application-specific questions have shown that users appreciate the adaptive composition and presentation functionalities and that the approach has delivered results in a form that matches a user state and query intent.

Despite certain differences appearing rather small, the statistical results from the paired t-tests have revealed significance with p-values of 0.03 or smaller across all metrics. This confirms that the results have not occurred by chance and that the adaptive composition approach has consistently performed high on user effectiveness, efficiency and satisfaction. Also, as users were unaware of the various tasks and systems before the experiment, there was no incentive for users to be biased towards either of the systems.

Many approaches have been suggested for evaluating adaptive systems. One approach consistently used in Adaptive Hypermedia publications has been holistic, which evaluates the overall effect of a system's adaptivity. Since the main target of the compositional approach presented in this thesis lies in an overall improvement of user assistance and guidance, a holistic task-based evaluation of the ARCHING prototype has been used in this chapter. By placing the user in the centre of the evaluation, real-life evidence has been provided for the benefits of adaptively retrieving and composing information presentations.

A common criticism of this approach is where the non-adaptive system is merely a version of the adaptive system with all adaptivity turned off. However, this was not the case in the presented study, since the non-adaptive system was a purpose-built search engine that simulated the conditions in a real corporate IR system typically used for the given tasks. It may be argued that the baseline system could be optimised in order to increase user effectiveness, efficiency and satisfaction. However, it has to be noted that the adaptive system is based on the exact same indexing and retrieval models and therefore an improvement of the baseline system would likely entail an improvement in the adaptive system as well.

While the evaluation in this chapter has provided clear evidence for the benefits of adaptive compositions and presentations, there exist a number of alternative prototype

possibilities for ARCHING. In particular, the adaptive strategies utilised in the PCC system only show one particular instantiation of the retrieval and composition capabilities provided by ARCHING. For example, in terms of open-corpus retrieval, the presented PCC system has only operated over a harvested cache and has not made use of the open-web retrieval functionalities. Moreover, as the compositional strategy has focused on complementing initial closed-corpus retrieval results with open-corpus forum entries, the classification modules have not been required in this example. Similarly, there are a number of other unexplored adaptation dimensions, such as a user's language capabilities. Also, there exist a number of alternative compositions in terms of result grouping and overall presentation.

Lastly, the evaluation in this chapter has focused on (i) homogeneous customer care tasks in terms of question types (i.e. each task contained a range of introductory, instructional and problem-solving questions) and (ii) a homogeneous user group in terms of search and domain expertise (i.e. all users were experienced search users and had no/little domain knowledge). In order to provide more fine-grained evaluation results, it is therefore necessary to investigate varying benefits of adaptive information compositions across (i) heterogeneous tasks and (ii) heterogeneous user groups.

4.6. Conclusions

Following the conclusions of the initial Adaptive Composition system evaluation in chapter 3, this chapter has presented a set of design principles for an extended architecture to overcome the identified limitations. In particular, the design principles include the retention of adaptive composition and presentation principles of the first iteration, while flexibly integrating free-text keyword search and open-corpus manipulation and adaptation capabilities.

An extended architecture called ARCHING (Adaptive Retrieval and Composition of Heterogeneous INformation sources for personalised hypertext Generation) has been presented, which flexibly integrates lightweight open-corpus adaptation functionalities into the adaptive composition and presentation architecture of the first iteration. A Personalised Customer Care (PCC) prototype implementation has demonstrated the maintained adaptive composition, presentation and navigation functionalities, as well as the successful integration and composition of large-scale open-corpus information. The

prototype has been evaluated in a real-life customer care case study, where users are asked to search for information relating to authentic information needs in order to complete a set of task questions as quickly as possible.

The evaluation results have revealed the benefits of the compositional approach to information retrieval and delivery, as it is shown to significantly enhance a user's efficiency, effectiveness and satisfaction compared to a conventional search system. As opposed to the initial adaptive composition system presented in chapter 3, ARCHING allows users to input conventional free-text queries, while still receiving an adaptively composed hypertext response. It is also shown that the open-corpus content can be successfully integrated into the adaptive composition without requiring any additional metadata. Furthermore, user satisfaction questionnaires have revealed that this integration is highly appreciated by users, stating that the compositions are easier to navigate and that they feel much more guided across the content sources. Moreover, users recognise and value the personalisation aspects in adaptive compositions, stating that the system provides guidance to personally relevant content.

Similar to the findings presented in chapter 3, it is shown that adaptive compositions encourage and motivate users to navigate across more content, resulting again in an increased page visit count. However, as opposed to the educational scenario, this effect appears to be weaker in the PCC prototype and therefore does not result in an increased task time. For the presented customer care scenario, this can be seen as a positive result, as users should get encouraged to "learn" as much as possible about the software product without suffering from a negative effect on their problem resolution time.

Overall, this chapter has revealed that the compositional approach can be applied successfully across heterogeneous information sources, including both structured (metadata-rich), as well as unstructured (metadata-sparse) content. It has been shown that the metadata-richness of closed-corpus information can successfully drive adaptive responses, and that open-corpus adaptation functionalities can fully integrate unstructured content into the generated compositions. Through the seamless integration of both closed-corpus and open-corpus content, users can adaptively navigate through structured and unstructured content, benefitting from both the quality of professionally authored content, as well as the quantity, diversity and freshness of open-corpus content.

Moreover, the presented results are not solely encouraging for information system end-users. In particular, from an information provider's perspective, developing professional content constitutes an expensive process and they would therefore like to maximise the usage of such information. The presented adaptation and personalisation capabilities can be seen as key enablers to successfully reuse such content by tailoring the delivery of information to particular user preferences, needs and context. While current corporate information systems have not been able to leverage the benefits of other material already on the web, the presented techniques and technologies can combine and compose the professionally authored content with open-corpus information that is generated by a company's own end-users and communities (e.g. in blogs, forums, etc.).

Lastly, while the PCC system presented in this chapter has provided clear benefits for the use of the compositional approach to information retrieval and delivery, there remains a multitude of alternative composition possibilities using ARCHING. In particular, alternative system prototypes may vary in terms of information source selections, result compositions or overall presentations. In addition, different composition types may suit particular user characteristics, as well as different task types. Chapter 5 explores such additional possibilities through the design, development and evaluation of multiple prototypes and examines the suitability of different interface compositions for different user needs and characteristics.

Moreover, there are a multitude of additional adaptation dimensions that can be supported using the presented architecture, such as multilingual capabilities or different device interfaces. A number of ARCHING implementations that highlight this additional dimension support are presented and evaluated in chapter 6.

5 Investigation & Evaluation of Open-Web Personalisation: A Comparative Approach

5.1. Introduction

Chapter 4 has introduced an architecture called ARCHING (Adaptive Retrieval and Composition of Heterogeneous INFORMATION sources for personalised hypertext Generation), which can adaptively retrieve, recombine and present closed-corpus and open-corpus information sources to support users in a personalised manner. This architecture has been successfully evaluated using an initial prototype implementation in the domain of customer care, providing evidence that the compositional approach to information retrieval and delivery can enhance a user's effectiveness, efficiency and satisfaction. However, this initial prototype implementation represents only one example of the capabilities of the presented architecture. There exists a multitude of additional retrieval and adaptation possibilities, which allow the generation of alternative interface compositions, as well as the integration of diverse open-web information.

First of all, this chapter investigates a number of alternative interface compositions and presents further evidence for the successful application of adaptive retrieval, composition and presentation. In particular, this chapter presents three distinct interface compositions developed using the ARCHING architecture (including the adaptation within these compositions) and investigates their respective suitability for different user needs and characteristics.

The second objective of this chapter lies in evaluating the retrieval, composition and presentation of *open-web* information. More specifically, while the initial prototype implementation presented in chapter 4 has operated over a harvested cache of open-corpus information (focussed on user forums), the underlying ARCHING architecture contains additional *open-web* capabilities that can retrieve information without prior harvesting (see section 4.3.2.1 for a description of the technical details). The prototypes presented in this chapter each make use of such capabilities and thereby dynamically retrieve and compose *open-web* information at run-time without prior harvesting of open-web content.

As the prototypes operate over such open-web information, an initial survey was conducted to investigate frequently used web sources in a customer support scenario (in order to inform the targeted open-web retrieval). The results from this survey are presented in section 5.2. Section 5.3 then presents a series of three distinct interface composition designs for this customer care scenario. This initial design phase also includes a small-scale user study, which was performed using paper-based interface design mockups. Section 5.4 presents the implementations of these prototype designs, including the different adaptation process steps as well as example screenshots. As mentioned before, each of these three implementations integrates the open-web retrieval capabilities provided by ARCHING, allowing the retrieval and composition of information from the open-web, as well as targeted websites identified in the survey. Section 5.5 presents a comparative task-based user evaluation and discusses the relative performances of each of the three prototypes. Results confirm that users generally appreciate the adaptive composition and presentation of heterogeneous information sources and that such open-web compositions can successfully help users in completing authentic real-life tasks. Moreover, it is shown that users exhibit varying search behaviours for different composition types and that their respective appreciation and satisfaction can depend on particular preferences and characteristics. Finally, section 5.6 concludes this chapter with a discussion of the overall findings.

5.2. Open-Web Information Source Selection

As mentioned in section 5.1, each of the prototypes described in this chapter composes and presents information that is retrieved using both the closed-corpus retrieval as well as the open-web retrieval capabilities of ARCHING. Since these open-web capabilities

include the possibility to adaptively focus the retrieval on particular website domains (see section 4.3.2.1 for implementational details), it is important to first gather a set of relevant information sources that are typically frequented by users for particular intents.

Since the prototypes presented in this chapter were aimed at customer support, a small-scale online survey was conducted, which asked users to indicate general information source preferences in customer care scenarios. In particular, they were asked to specify their website preferences for the following query intents⁴⁶ (see APPENDIX C for full questionnaire).

(a) Introductory/overview information for product features

(b) Instructions/how-to information for product features

(c) Solutions to problems with a product

A total of 37 participants took part in this survey, revealing noticeable differences in preference ranks for varying information intents. Table 5-1 presents the aggregated preference ranks for the given intents, showing that for the respective questions, users mostly prefer product manuals for (a), support articles on the company website for (b) and forums for (c).

Additionally, when asked about which other websites they used for product support, participants most frequently mentioned Wikipedia.org for (a) (5 participants) and stackoverflow.com for (b) and (c) (9 participants).

It is worth noting that the presented questionnaire questions were generic in nature and that they were not focussed on Symantec products or services. It is arguable that the findings from the survey could therefore be reused for different customer support applications (e.g. for Microsoft Office support). However, in order to apply the approach to a different application area (e.g. elearning, cultural heritage), it would be necessary to perform a new survey to identify generic types of websites for this area. The results from that survey could then be used in the same manner to inform the targeted open-web retrieval components.

⁴⁶ These intents had been identified through an analysis of the support data provided by Symantec. As mentioned in chapter 4, this data contained real-life customer support logs, including exact queries submitted by customers.

(a) Where would you look for introductory/overview information for product features? (e.g. what is feature X responsible for?)	Aggregate Rank
Product manual / Built-in help	1
Support articles on company website	2
Forums	3
Other websites (e.g. ehow.com)	4
(b) Where would you look for instructions/how-to information? (e.g. how do I configure the proxy settings for this product?)	Aggregate Rank
Product manual / Built-in help	2
Support articles on company website	1
Forums	3
Other websites (e.g. ehow.com)	4
(c) Where would you look for solutions to problems with the product? (e.g. what should I do when error message XYZ appears?)	Aggregate Rank
Product manual / Built-in help	3
Support articles on company website	2
Forums	1
Other websites (e.g. ehow.com)	4

Table 5-1. Information Source Preferences

5.3. Design

Following this initial survey, a set of three distinct interface composition designs was developed, which varied significantly in terms of their interface structure and information coherence. In order to gain initial user feedback for these interface designs (i.e. before developing the respective implementations), paper-based mockups were developed and presented in a small user study (6 participants).

In this study, each user was presented with all three paper-based mockups. Moreover, the mockups were shown in random order. For each paper mockup, participants were requested to indicate their initial reaction to the interface using “reaction cards” (Benedek and Miner, 2002) (see APPENDIX D). In this method, participants are asked to choose a set of terms (from a fixed list of terms) that most closely correspond to their impressions of the presented interface mockup. Participants are then asked to elaborate on why they have chosen the particular terms and also where they might foresee any usability issues.

Sections 5.3.1, 5.3.2 and 5.3.3 provide brief descriptions of the different interface composition mockup designs, their respective adaptation possibilities, as well as a discussion of the initial user feedback. Each of the presented interface compositions is built around the notion of supporting a customer according to a query intent, which can consist of either (a) *I want to find out the basics*, (b) *I want to get a how-to* or (c) *I want to solve a problem*. Section 5.3.4 concludes the evaluation of mockups and provides a discussion of the overall mockup design findings.

5.3.1. Design of Interface Composition 1: Information Source Panels

The first composition presents a “panel-based” interface, whereby information is grouped into separate panels based on the underlying information source and/or information type. Figure 5-1 shows the presented mockup of such a composition.

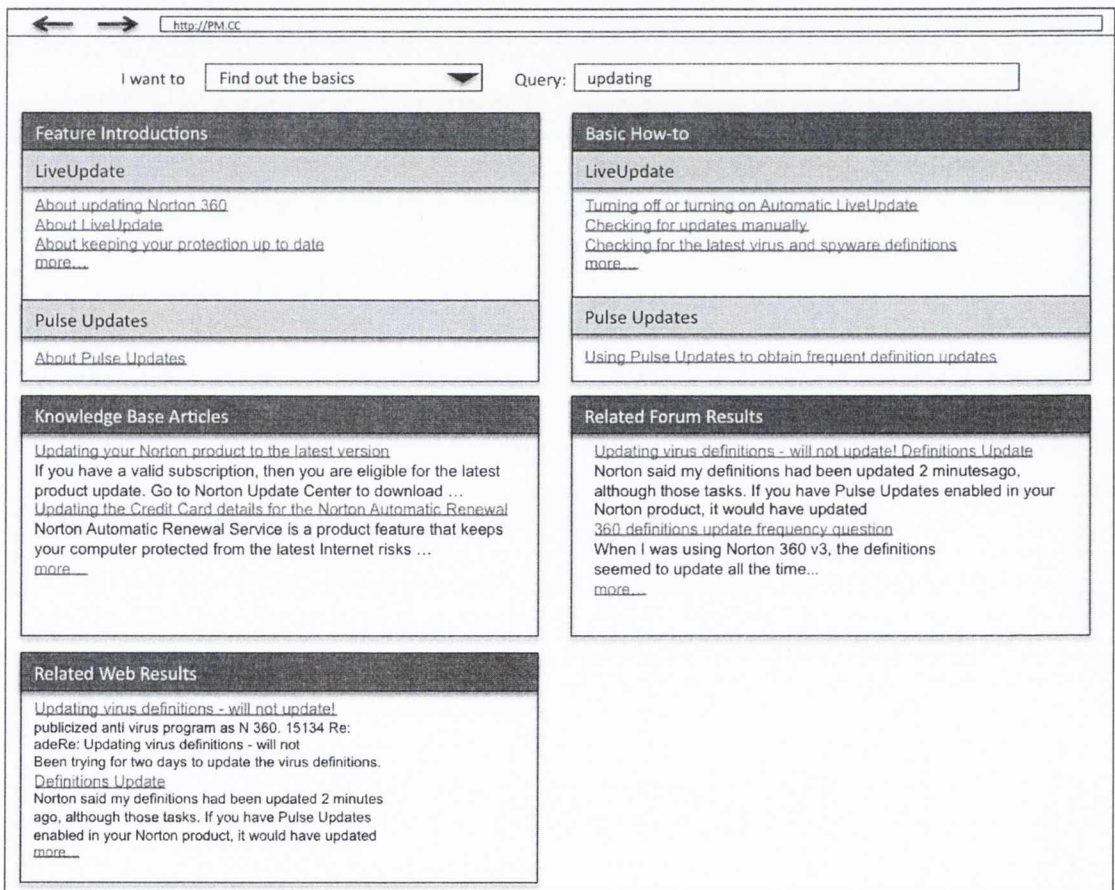


Figure 5-1. Composition 1 Mockup

In this figure, the top two panels contain information from the product documentation, with explanations being presented on the left and instructions on the right. Also, within these top two panels, information is further grouped according to product features (e.g.

“LiveUpdate”, “Pulse Updates”). The remaining panels contain information from the Norton Support website (“Knowledge Base Articles”), the Norton Community Forums (“Related Forum Results”) and the open web (“Related Web Results”).

The most significant adaptation potential with this type of composition lies in the adaptive positioning of panels according to the indicated query intent. Moreover, panels may be expanded/collapsed automatically, as well as altered in terms of their size and colour. In addition, the information within closed-corpus-based panels may be reordered depending on the most suitable metadata values.

When presented with this paper-based mockup, participants’ reaction card choices indicated that the interface generally felt “busy” (4/6), “organized” (3/6) and “time-consuming” (3/6) (Table 5-2 presents all the terms chosen by participants for this interface). In particular, participants felt that this composition may require significant effort to glance at multiple panels in order to get an overview of the results. However, participants also acknowledged that the panels represented a clear way of organising multiple information sources.

In terms of usability issues, participants noted that the naming of the panels was critical to the successful application of this interface. It was therefore recommended that the respective information sources should be named more clearly, e.g. “Norton Community Forums” instead of “Related Forum Results”. Moreover, while they acknowledged the potential benefit of rearranging panels (and altering their size/colour) according to different query intents, participants noted that this could result in an increased cognitive effort. However, participants also stated that this effort could decrease with continued usage of the system.

Card	Responses	Card	Responses
Busy	4	Disruptive	1
Organized	3	Distracting	1
Time-consuming	3	Dull	1
Clear	2	Easy to use	1
Consistent	2	Frustrating	1
Disconnected	2	Helpful	1
Overwhelming	2	Straight Forward	1
Integrated	1	Useful	1
Comprehensive	1	Valuable	1
Customizable	1		

Table 5-2. User reaction cards for Interface Composition 1 Design Mockup

5.3.2. Design of Interface Composition 2: Topic-based Composition (Highly structured)

The second composition is based on the initial prototype presented in chapter 4, which consists of a two-stage navigation process. The first stage presents an overview of initial results and groups the information according to the respective topics (i.e. product features) (see Figure 5-2). For each topic (e.g. Live Update), information is retrieved from the product documentation (left), the Norton Support website (middle) and the Norton Community forums (right).

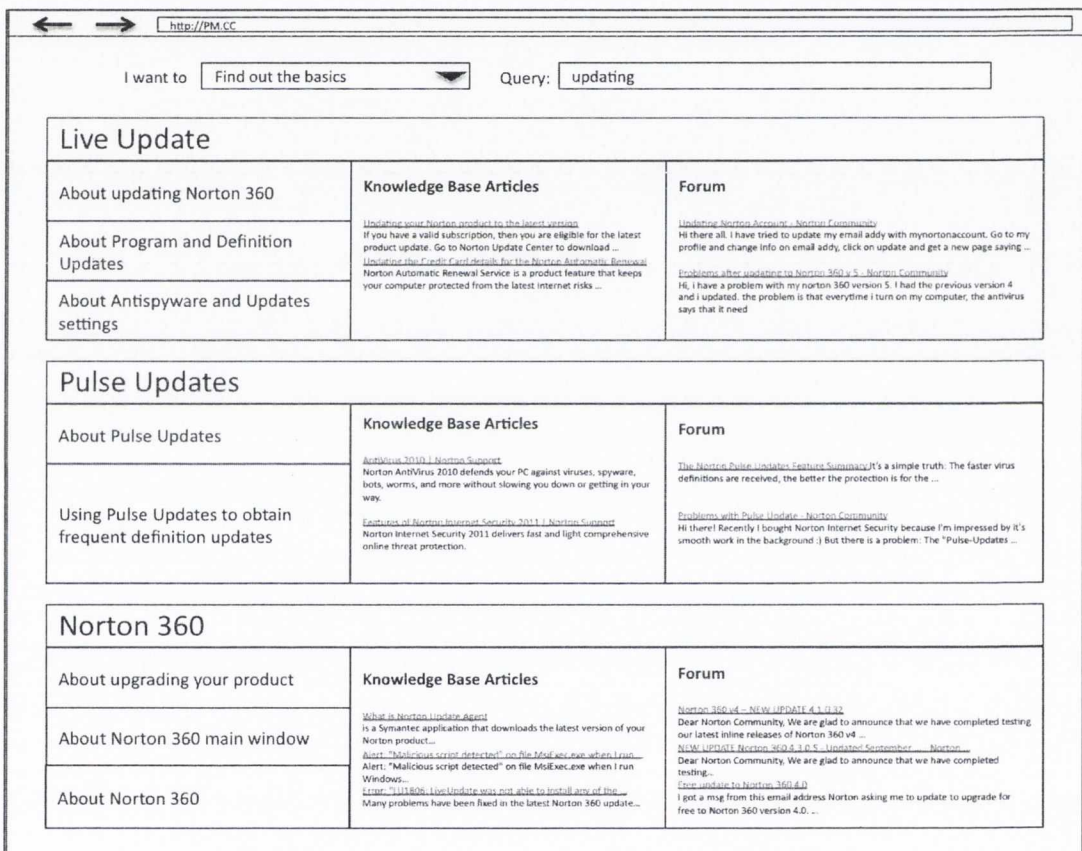


Figure 5-2. Composition 2a Mockup

After selecting one of the initial results from the first screen, the result content is displayed along with the second stage of the composition (see Figure 5-3). This second screen presents a highly structured overview of the topic related to the selected result. On the left, a tree-based navigation groups closed-corporus result titles according to their metadata values. In addition, the various information sources can be accessed through a series of tabs (i.e. “Manual“, “KBA“, “Forum“, “Web”). By using this double navigation, a user can choose to either directly access product documentation content through the links on the left (while having the “Manual” tab selected), or retrieve

focussed open-corpus result lists using the remaining tabs. These open-corpus result lists are generated based on the initial user query as well as the currently selected documentation topic.

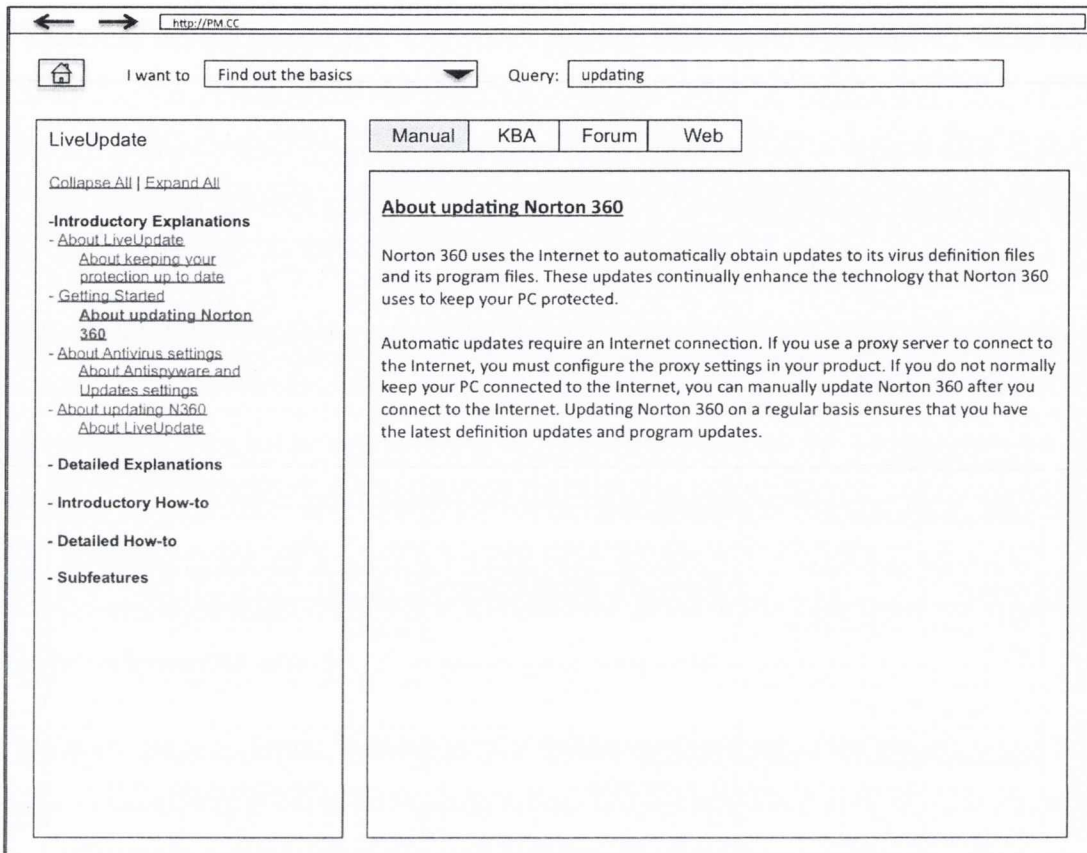


Figure 5-3. Composition 2b Mockup

As demonstrated by the prototype presented in chapter 4, there are a number of adaptation possibilities for this type of composition. First of all, on the overview result screen (Figure 5-2) closed-corpus results can be reranked based on the metadata values that most closely match the user's query intent. This also enables open-corpus adaptation, since open-web results are retrieved based on these reranked results. Moreover, the selection of initial open-corpus information sources can be adapted to the particular query intent. Secondly, on the structured result screen (Figure 5-3) the tree-based navigation can be reordered to promote the most suitable type of information (e.g. explanations, how-to). In addition, adaptation possibilities include the automatic expansion/collapsing of relevant/irrelevant information. For example, in the presented mockup interface the "Introductory Explanations" are automatically expanded as this corresponds most closely to the chosen query intent (i.e. *I want to find out the basics*).

Lastly, the information source tabs can be reordered in order to promote relevant sources for the chosen intent (e.g. forums for problem solutions).

When presented with this paper-based mockup, participants most notably mentioned that the composition was “Organized” (5/6) and “Understandable” (3/6) (Table 5-3 presents all the terms chosen by participants for this interface).

Card	Responses	Card	Responses
Organized	5	Convenient	1
Understandable	3	Efficient	1
Comprehensive	2	Effortless	1
Easy to use	2	Innovative	1
Helpful	2	Integrated	1
Meaningful	2	Simplistic	1
Clean	1	Straight Forward	1
Clear	1	Useful	1
Connected	1	Valuable	1

Table 5-3. User reaction cards for Interface Composition 2 Design Mockup

In particular, it was generally acknowledged that such a composition would allow an easy and comprehensive exploration of results across a number of distinct information sources. Moreover, participants generally felt that this interface presented a more integrated composition of results.

In terms of usability issues, users noted that the system should better highlight the fact that results were based on an expanded version of the original user query. This is particularly important on the second screen, where users select one of the open-web result tabs and then use the links on the left to explicitly modify the original user query. Moreover, users generally felt that a reordering of the tree-based navigation was not necessary and that it might even result in a more confusing interface (expanding/collapsing was deemed sufficient). Participants also tended to oppose the reordering of information source tabs, as this might again result in a confusing interface.

5.3.3. Design of Interface Composition 3: Topic-based Composition (Loosely structured)

The third composition combines ideas from both Composition 1 and Composition 2. First of all, this composition consists again of a two-stage process, whereby the first screen presents an overview of information that is grouped according to topics (i.e. product features) (see Figure 5-4).

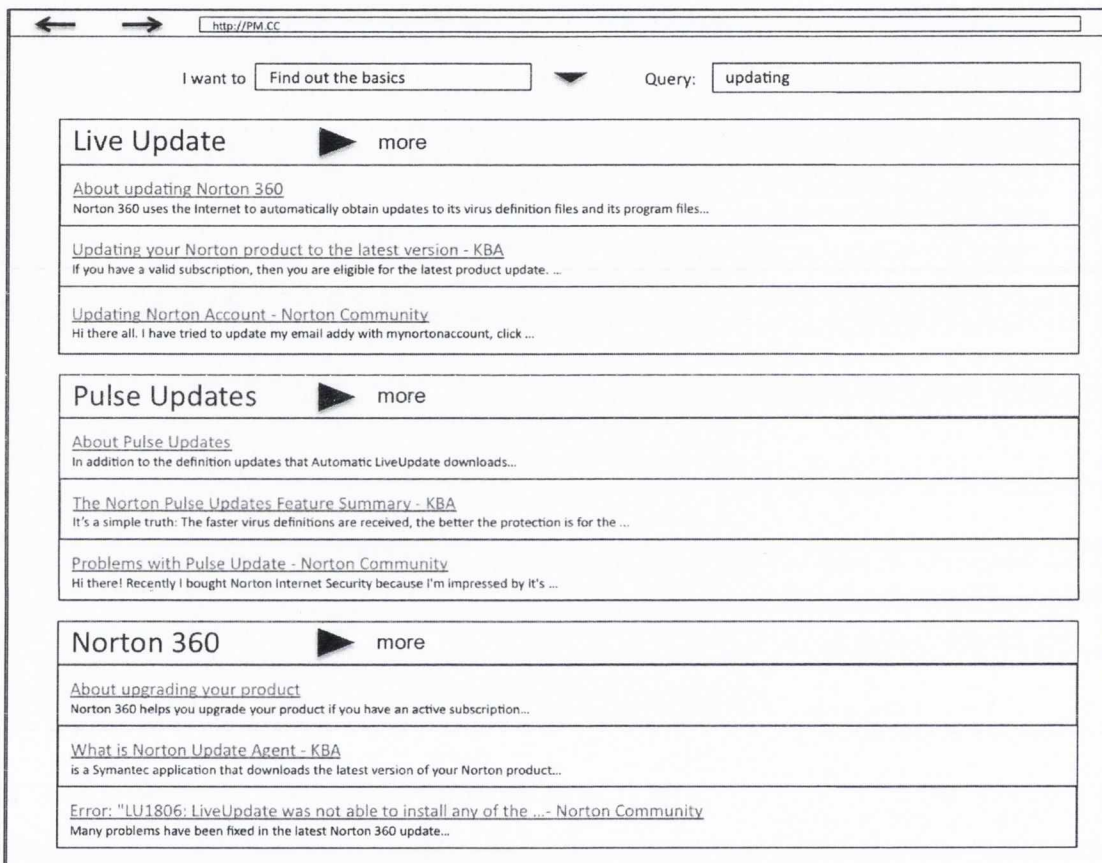


Figure 5-4. Composition 3a Mockup

For each topic, information is again retrieved from a number of information sources, although there is no strict information source separation as in Composition 2 (the results are only presented in a flat ranked list for each topic). If users cannot fully satisfy their information need, it is possible to get a more comprehensive information composition for each topic. By clicking on the respective “▶ more” button, users are presented with an interface that resembles more closely Composition 1. This interface clearly separates the various information sources into distinct areas of the screen. For example, Figure 5-5 shows how this interface presents results related to the “Live Update” feature from the product documentation (“Manual”), as well as the open-web (“Web Results”). Moreover, information that is semantically related to the currently selected

product feature is also displayed alongside (“See also”). As shown in Figure 5-5, the comprehensive composition also generates a grouping according to the type of information provided by the respective sources (only one grouping is shown at a time). For example, the “About” grouping contains introductory explanations, “Instructions” contains how-to information and “Problem solutions” contains content that addresses particular product issues.

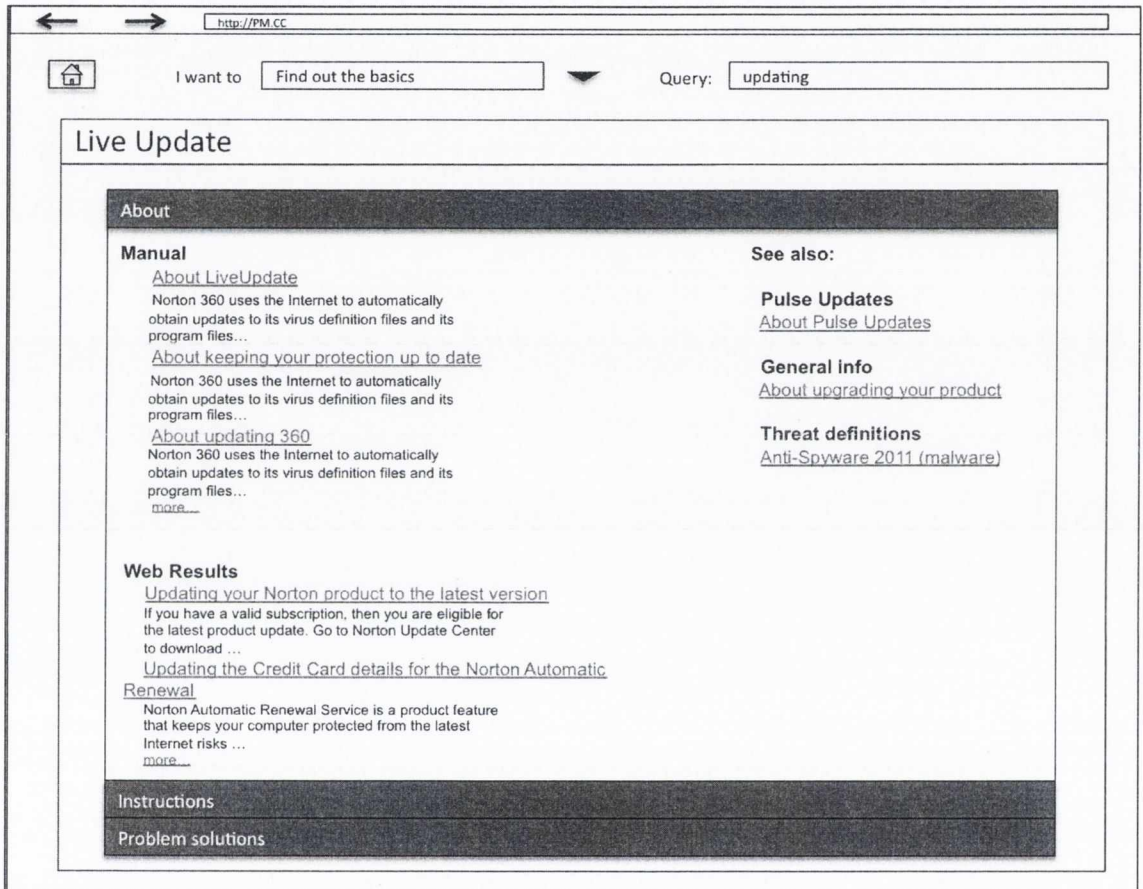


Figure 5-5. Composition 3b Mockup

Similar to Composition 1 and Composition 2, there are a number of adaptation possibilities for this composition. First of all, on the first screen it is possible to adaptively select the most appropriate information sources for the indicated query intent. Moreover, closed-corpus results can be reranked based on the metadata values that most closely match this intent. On the second screen, adaptation can occur in a similar fashion to Composition 1, by adaptively positioning or expanding/collapsing the various information source sections.

In terms of user reaction, participants tended to feel that this composition mockup was “Organized” (5/6), “Straight forward” (3/6) and “Calm” (3/6) (Table 5-4 presents all the terms chosen for this interface).

Card	Responses	Card	Responses
Organized	5	Integrated	2
Calm	3	Meaningful	2
Straight Forward	3	Overwhelming	2
Appealing	2	Understandable	2
Clean	2	Connected	1
Clear	2	Disruptive	1
Consistent	2	Efficient	1
Controllable	2		

Table 5-4. User reaction cards for Interface Composition 3 Design Mockup

It was generally acknowledged that the two-stage presentation could be beneficial in terms of serving single-answer queries (using the first screen), as well as more complex questions that require the synthesis from a number of sources.

However, participants also pointed out a number of potential usability issues of such a composition. First of all, on the initial overview screen, participants requested to make better use of the space available on a display. It was generally advised that the composition should more closely follow the principle of Composition 2, by providing a slightly more structured display of various information sources. In addition, participants noted that the additional grouping of “About”, “Instructions” and “Problem solutions” in the second screen might be cumbersome to use and that the display should therefore more closely follow the simple principles of Composition 1.

5.3.4. Discussion

Overall, the evaluation of the paper-based interface design mockups has provided initial evidence for the strengths and weaknesses of different composition types.

First of all, participants have generally responded positively towards the notion of adaptively selecting, composing and presenting results from a number of distinct information sources (as the majority of reaction cards selected by the participants were of a positive nature). In particular, participants noted that the results in each of the compositions were very organised (especially in compositions 2 and 3). However, due to the less integrated nature of composition 1, participants indicated that such an interface generally feels busier and therefore might be more time-consuming than the topic-structured compositions 2 and 3. Similarly, compositions 2 and 3 were generally regarded as being more comprehensive and understandable.

In terms of adaptation capabilities, the compositions present a number of possibilities for tailoring results towards the different user intents. While composition 1 mainly relies on the automatic reordering and resizing of the various panels, there are several additional adaptation possibilities for compositions 2 and 3, for example using the closed-corpus metadata to appropriately structure and sequence the product documentation as well as the related open-web results. However, throughout the discussion with the participants, it was also noted that it is crucial for the compositions not to over personalise the navigation structure too much, as this might result in confusing and inconsistent interfaces.

Lastly, by running this initial evaluation of mockups, it was also possible to identify a number of potential usability issues early on. Most importantly, participants requested a clearer naming of the various information groupings in order to allow users to recognise the exact provenance of the results. For example, terms such as “Knowledge Base Articles” were deemed too technical and should be replaced with more meaningful labels (e.g. “Norton Support Articles”). Similarly, it was regarded as crucial to clearly communicate to the user which query was responsible for the retrieval of the presented results (e.g. original vs. expanded). Moreover, in some cases (in particular composition 3), participants reacted negatively towards navigation structures that were too structured (hence requiring too many navigation steps) and it was recommended to slightly simplify the interface design.

5.4. Implementing the interface compositions

Each of the composition designs described in this section has been implemented using the ARCHING architecture (presented in chapter 4). The major difference between the three prototypes lies in the way in which the various architecture functionalities are selected and combined by the adaptation narrative in order to produce the respective composition presentations. Sections 5.4.1, 5.4.2 and 5.4.3 describe the particular adaptation processes used by these implementations and illustrate the respective compositions with screenshots of the resulting user interfaces.

5.4.1. Interface Composition 1 Implementation (C1)

As outlined in section 5.3.1, the first composition presents a panel-based interface, where information sources are clearly separated from each other in distinct panels. Figure 5-6 presents the overall adaptation process for generating this composition.

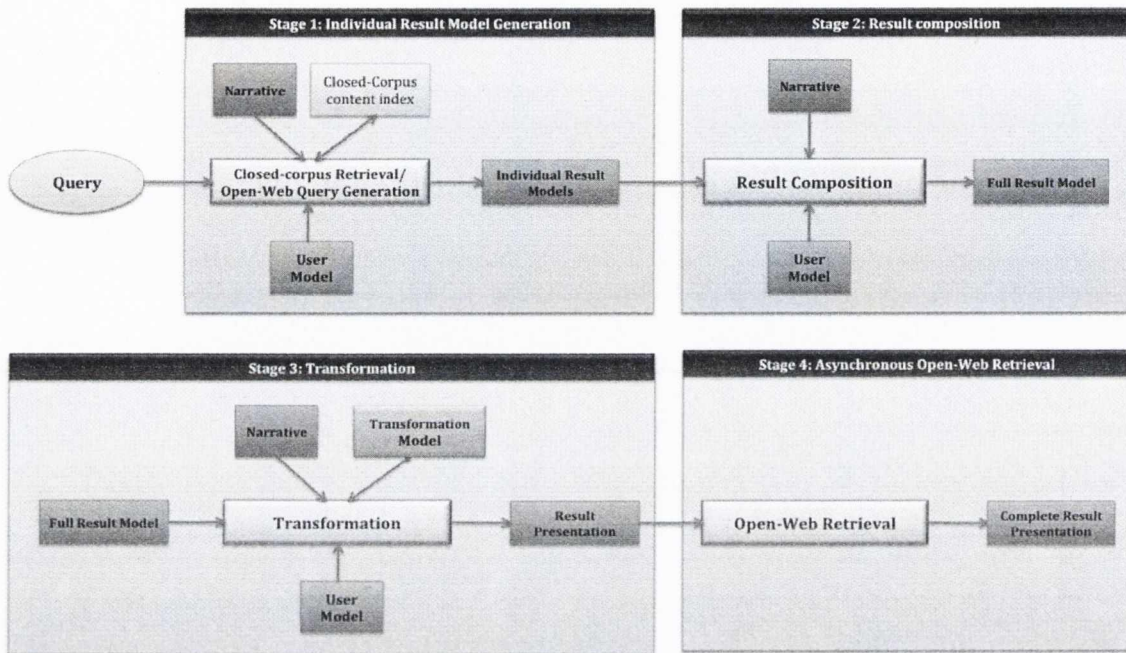


Figure 5-6. Composition 1 Adaptation Process

In the first stage, the user query is executed on the closed-corpus sources in order to retrieve an initial set of results. These results are also grouped according to their activity type in order to best match the user’s query intent at the composition stage (e.g. explanations are prioritised for “find out the basics” and instructions for “how-to”). In case there are no closed-corpus results retrieved (due to the query keywords not appearing in the closed corpus, e.g. in the case of a user inputting a specific error code), a temporary open-web search is conducted using the Bing API. The temporary results from this web search are classified to find the most related ontology classes (i.e. product features), for which closed-corpus results can then be retrieved using SPARQL queries.

In addition to the retrieval of closed-corpus results, skeleton result models are generated for the various open-web information sources. These models each contain a number of parameters, including the initial user query, as well as additional retrieval specifications such as a focused site domain, number of results, offset, source or language. While it is

possible to also execute these open-web queries at this stage, it is preferred to use the asynchronous retrieval capabilities (during stage 4) for efficiency reasons.

During the second stage, the closed-corpus result model and open-web skeleton models are composed according to the suitability for the chosen query intent (based on the survey results presented in section 5.2). For “find out the basics”, explanations from the product documentation are deemed most relevant, followed by Norton Support Articles, Norton Community Forums and finally Web Results. Conversely, for “How-to” queries, the prioritised sources are Norton Support Articles as well as instructions from the product documentation. For “problem solutions”, the most important sources are Norton Community Forums and Norton Support Articles. The composition is achieved by creating a “Full Result Model”, which groups the individual result models into a table-like structure (i.e. rows and columns) according to the prioritised sources (the most important sources are placed in the top row of this structure).

In the third stage, the full result model (XML-based) is transformed into XHTML, which can then be displayed to the user (see Figure 5-7 for a screenshot of the final implementation).

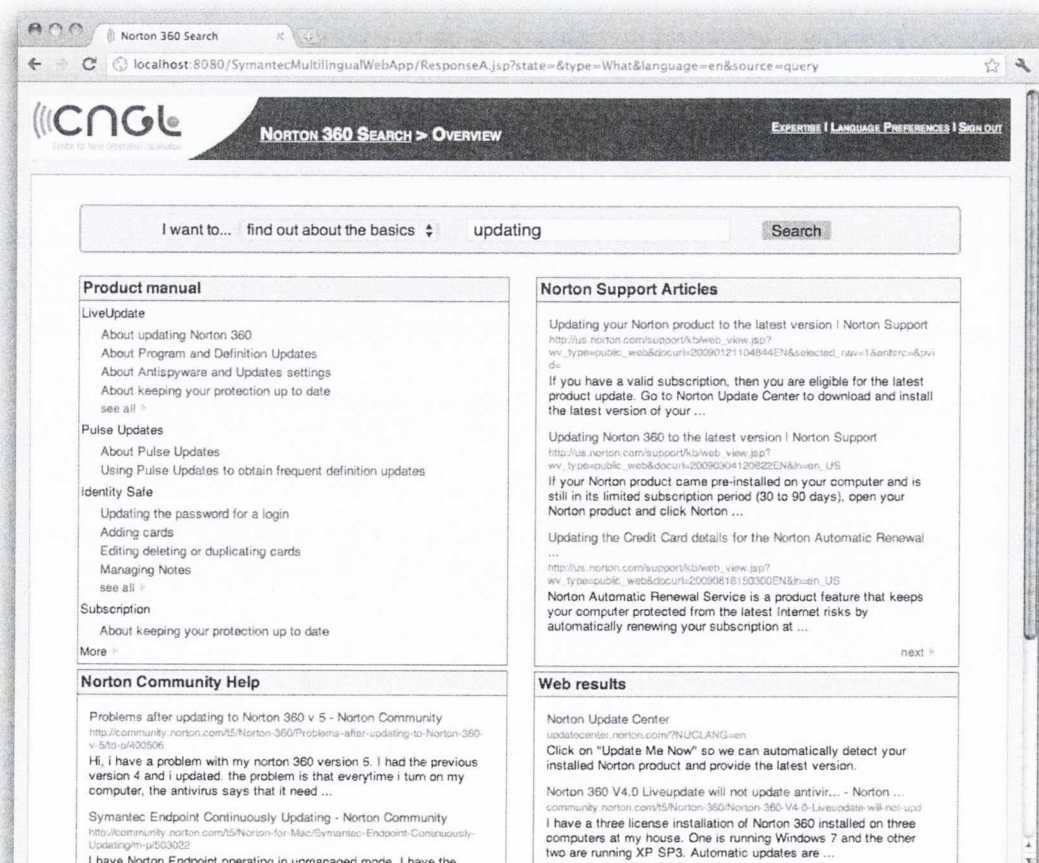


Figure 5-7. Composition 1 Screenshot

In the final stage, asynchronous (Bing) requests are executed using the open-web queries that were pre-specified in stage 1. The results of these requests are dynamically integrated into the composition structure in order to complete the final result presentation. The rules for this transformation are held in a dedicated model, which can be adaptively changed to suit different contextual characteristics such as device specifications (see section 6.6.1 for an example of mobile device presentation generation).

5.4.2. Interface Composition 2 Implementation (C2)

The second composition is based on the adaptation process of the initial customer care prototype described in section 4.4.3. In particular, this composition is generated through a number of adaptation steps, which group closed-corpus information according to associated topics (i.e. product features) and relate open-web information through query expansion. In addition to this process described in section 4.4.3, Composition 2 also draws from the various information sources that are used in Composition 1. These information sources are again retrieved asynchronously for efficiency reasons.

Moreover, in order to provide better coverage for unknown keywords that might occur for the query intent “solving a problem” (e.g. error codes that are not mentioned in the closed corpus), the adaptation process includes an extra step to first expand the original user query with keywords from an initial open-web search. More specifically, the results from a temporary open-web search (which executes the original query to the Bing API) are used to generate a list of frequently occurring terms, which are then used to expand the original query.

Figure 5-8 presents a screenshot of the final prototype interface (second screen), showing the various information sources that are available to a user through a tabbed interface. In this particular screenshot, a user has selected one of the introductory results from the product documentation (“About updating Norton 360”). Other sources that are available are “Norton Support”, “Forum” and “Web”.

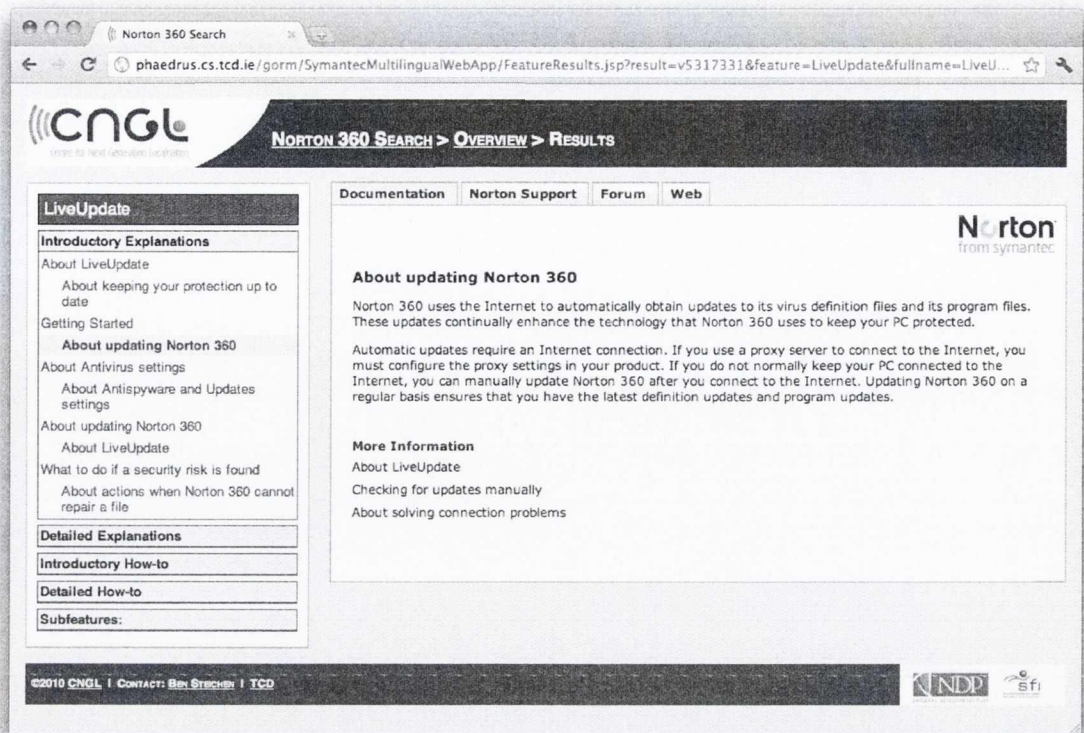


Figure 5-8. Composition 2 Screenshot (second screen)

5.4.3. Interface Composition 3 Implementation (C3)

As mentioned in section 5.3.3, the third composition (C3) constitutes a combination of ideas from C1 and C2. Similar to C2, it consists of an adaptive presentation that first presents an overview screen, where information is grouped according to topics (i.e. product features). In order to generate this overview screen, the same adaptation process is used as the initial prototype system presented in section 4.4.3 (see Figure 4-7 in particular). A user can also go to a second screen, which provides a similar composition as presented in Composition 1.

This (more comprehensive) composition presents the closed-corpus results related to the chosen topic, as well as open-web results that have been retrieved using an expanded query based on the closed-corpus result titles. In order to generate this second screen, a similar adaptation process is used as presented in 5.4.1 (see Figure 5-6 in particular). The only major difference lies in the fact that the user query is first expanded using the closed-corpus result titles for the selected feature in order to focus the open-web retrieval.

Figure 5-9 presents a screenshot of the second screen of C3, where a user has selected to view the comprehensive set of results related to the “LiveUpdate” feature. Similar to

the screenshot shown for C1 (see Figure 5-7), the various information sources are clearly separated from each other.

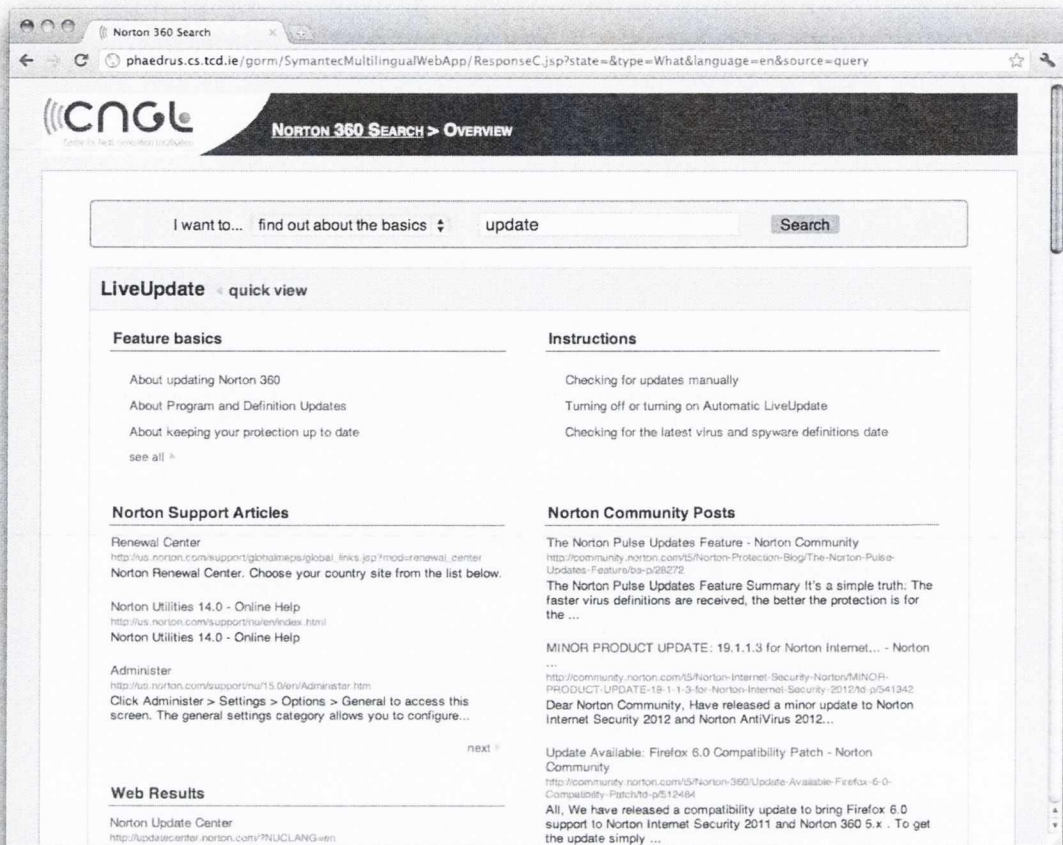


Figure 5-9. Composition 3 Screenshot (second screen)

5.5. Evaluation of interface compositions

In order to evaluate the presented compositions, a real-life user-study was performed using authentic information needs in the context of customer support (similar to the initial prototype evaluation presented in section 4.5). The main goal of this evaluation was to investigate the varying degrees of task assistance and user satisfaction for different composition types. In particular, the three prototypes were evaluated in terms of their ability to support users in real-life customer support tasks (user efficiency and effectiveness), as well the usability from the users' perspective (i.e. user satisfaction). Moreover, these aspects were evaluated for varying degrees of task difficulty and user characteristics in order to investigate more fine-grained differences between compositions. Lastly, since each of the composition prototypes made use of the open-web capabilities of the ARCHING architecture, the evaluation also investigated the overall usability of the open-web information integration.

5.5.1. Hypotheses/Sub hypotheses

Task Assistance

Similar to the evaluation criteria presented in section 4.5.1, the benefit to the user in a customer support scenario lies in a system's ability to assist a user's search for information effectively and efficiently. In particular, a composition system should require users to invest the least amount of effort in order to find as much relevant information as quickly as possible in order to complete their task. The hypotheses regarding the user effectiveness and efficiency were as follows.

- H1: The adaptive composition types provide different degrees of task assistance.
 - H1.1: The adaptive composition types provide different degrees of task assistance in terms of user effort for task completion.

The metrics used to test this hypothesis are completion time and number of queries issued.

- H1.2: The adaptive composition types provide different degrees of task assistance in terms of the amounts of relevant information viewed by users.

The metric used to test this hypothesis is the users' overall page view count.

- H1.3: The adaptive composition types provide different degrees of task assistance in terms of task completion effectiveness and perception.

The metric used to test this hypothesis is the users' measured and perceived task accuracy.

- H1.4: The adaptive composition types provide different degrees of perceived overall task assistance.

In order to test this hypothesis, usability questionnaire scores are compared across compositions.

User Satisfaction

In addition to the assessment of task assistance, the second goal of the evaluation was to measure the degrees of user satisfaction for different composition types. The hypotheses regarding user satisfaction are therefore as follows.

- H2: The adaptive composition types provide different degrees of user satisfaction.
 - H2.1: Overall, the adaptive composition types provide different degrees of usability.

Usability questionnaire scores are used to test this hypothesis.

- H2.2.: The adaptive composition types provide different degrees of usability for users with different characteristics.

In order to test this hypothesis, usability questionnaire scores are correlated with user characteristics captured during prequestionnaires.

- H2.3: Users recognise and value different aspects of composition, adaptation and personalisation for different compositions.

In order to test this hypothesis, application-specific usability questionnaire scores are compared across compositions.

Open-Web Information Integration

In addition to the comparative evaluations of the different interface compositions (H1 and H2), the third goal was to investigate the general usability of information compositions that integrate *open-web information*. To this end, the prototypes presented in this chapter each used the *open-web* retrieval capabilities of the ARCHING architecture, as opposed to retrieving open-corpus information from a harvested cache (such as the prototype presented in chapter 4). The overall evaluation results for the three prototypes thus captured the general usability of *open-web compositions*. In particular, this evaluation concentrated on the *overall task assistance and user satisfaction of the open-web prototypes*⁴⁷.

⁴⁷ Note that since the evaluation consisted of new task questions and varying task difficulties, it would not be meaningful to compare individual evaluation results (e.g. time on task, number of queries) directly to the open-corpus prototype evaluation presented in chapter 4.

The corresponding hypothesis was as follows:

- H3: The open-web composition and integration with the closed-corpus content provides (on average) positive results for task assistance and user satisfaction.

In order to test this hypothesis, aggregate results for all three prototypes are calculated from the metrics presented for H1 and H2.

5.5.2. Experimental Setup

The experimental setup followed a similar methodology to the initial prototype evaluation presented in section 4.5.1. A set of 3 tasks was developed for the user study, which consisted again of real-life information needs regarding the Symantec product Norton 360 (taken from Symantec training material and customer support interaction topics). Each task contained a set of 4 to 5 questions regarding various aspects of the product. Moreover, the three tasks were designed to be of varying difficulty in order to evaluate the relative assistance of the three composition prototypes. In contrast to the evaluation presented in section 4.5.1, each user received only one of the composition systems for the completion of all three tasks. This ensured that for each prototype, tasks 1, 2 and 3 were completed an equal number of times. (The exact task questions, as well as the questionnaires can be found in APPENDIX E.)

The process for each user started by receiving an e-mail about the purpose and length of the experiment, as well as the experiment URL and personal credentials. After logging in to the experiment system, users were first asked to fill out a consent form.

In order to capture various user characteristics and preferences, users were then presented with a series of prequestionnaire questions (to allow the testing of hypothesis H2.2). First of all, users were asked to indicate their *background regarding the Symantec product Norton 360* (e.g. their experience and expertise), as well as their *general experience with search systems* (e.g. their knowledge of advanced search features). Secondly, users were asked about particular *characteristics regarding knowledge acquisition and cognitive preferences* in customer support scenarios (e.g. sequential vs global learning). Thirdly, as the various composition prototypes had significantly different interface designs, *general search user interface preferences* were captured to analyse possible correlations. In order to capture such general user interface preferences, users were presented with a series of four website screenshots and a

number of accompanying questionnaire questions. These screenshots represented significantly different types of information access interfaces, ranging from a panel-based interface with low information coherence to a fully-structured result presentation with highly integrated information. Table 5-5 shows the overall characteristics of the different website screenshots shown to the user, each of which presented search results on the same topic. The full questionnaire and website screenshots can be found in APPENDIX E.3.

	Screenshot 1: WDYL⁴⁸	Screenshot 2: Yippy⁴⁹	Screenshot 3: HowStuffWorks⁵⁰ (1)	Screenshot 4: HowStuffWorks⁵¹ (2)
Interface structure	Panel-based information source separation	Ranked list, clustered navigation	Ranked list, semi-structured results	Fully-structured result presentation
Coherence of information	low	medium	medium	high

Table 5-5. Prequestionnaire website screenshots, used to capture general search interface preferences

Users then received instructions on how to use their given composition system, including a short video tutorial and the chance to test the system using a test task. After users felt confident with the system functionalities, they would then proceed to their first task screen (displayed as task A). This screen consisted of a set of questions, which users were asked to answer using the provided text boxes. In order to solve the task, users were given a link to the composition system to allow searching across the various content sources. After the completion of their first task, users were asked to fill out a short questionnaire regarding the perceived difficulty of the task, their perceived task performance as well as their satisfaction with the system assistance and guidance (using a Likert scale ranging from 1=strongly disagree to 5=strongly agree). Following this, users proceeded to perform their second and third tasks (displayed as task B and C respectively) in the same manner, each followed by the short questionnaire described above.

⁴⁸ wdy.com

⁴⁹ yippy.com

⁵⁰ Howstuffworks.com

⁵¹ Howstuffworks.com

Users were then asked to fill out a SUS questionnaire (Brooke, 1996), as well as an application-specific usability questionnaire (again using a 5-point Likert scale). After completing the full evaluation process (see Figure 5-10), users were automatically entered into a random draw for the chance to win an electronic device.

The experimental process was entirely online and users were asked to perform the experiment in a single session without interruption. User actions were tracked throughout search sessions in order to be able to analyse users' system interaction behaviour. Also, task completion times were tracked between the first display of a task's questions until a user's submission of the task answers. In order to balance any effects of order bias, tasks were assigned using Latin square design.

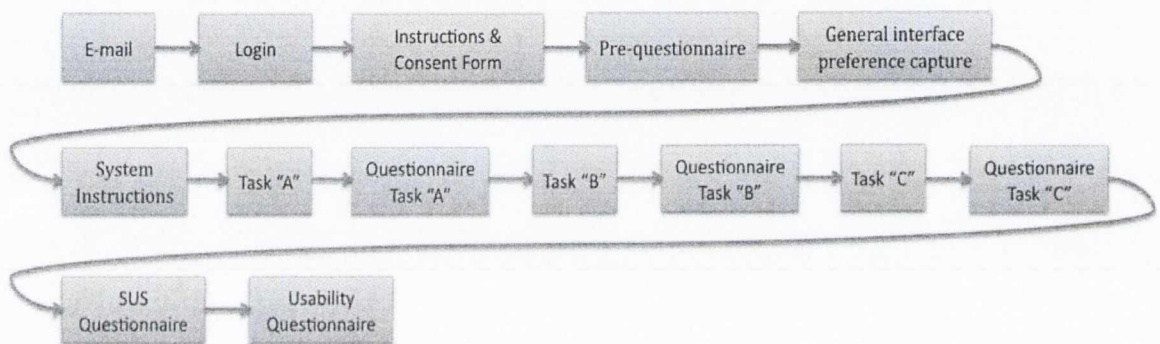


Figure 5-10. Experimental process

5.5.3. Results

A total of 158 users signed up for the online study, out of which 87 completed the full evaluation process. Participants were recruited from Trinity College Dublin, the University of South Australia, the University of Victoria (Canada), as well as Symantec Corporation.

User characteristics

The pre-questionnaire revealed that all users were daily computer, web and search system users. The additional user characteristics gathered from the prequestionnaire are shown in Table 5-6. As can be seen from this table, there are a number of differences among the various users. This information will be used below to determine user satisfaction correlations between specific user characteristics and the different composition systems.

Also, there were no significant differences between the various system groups (see APPENDIX F for detailed prequestionnaire results).

Would you be able to advise people on using and configuring Norton 360?*	
Yes	63
No	24
Have you ever used advanced search engine features (e.g. using the '-' sign to specify unwanted terms)?**	
Yes	60
No	27
What do you tend to do when encountering a software problem?	
Self-help (through manuals, forums, web searches, etc.)	81
Contact the help/call centre	6
Which of the following statements applies to you most?	
I like getting a quick how-to/fix without additional explanations.	47
I like understanding the cause of a problem that has occurred.	40
What is more important to you?	
A webpage lays out the content in clear sequential steps.	72
A webpage gives me an overall picture and relates the content to other subjects.	15
How would you generally understand new software features?	
Once I understand all the parts, I understand the whole thing.	54
Once I understand the whole thing, I see how the parts fit.	33

*. this question was aimed at estimating a user's domain expertise

** . this question was aimed at estimating a user's search expertise

Table 5-6. Overall prequestionnaire answers

Overall Task Characteristics

As mentioned previously, tasks were designed to be of varying difficulty in order to investigate more fine-grained differences between the compositions systems (task 1 being the easiest and task 3 being the hardest). When analysing the combined measured data across all systems, it was confirmed that there were significant differences between

the tasks in terms of perceived complexity ratings, average time on task and average number of queries.

Table 5-7 presents an overview of these overall task characteristics. As can be seen in the top sub-table (*Average Perceived Complexity*), the average perceived complexity was lowest for task 1 (average user rating of 2.74 out of 5), and highest for task 3. The difference between task 1 and task 2 (T1,T2), as well as the difference between task 1 and task 3 (T1,T3) were also found to be statistically significant using paired t-tests ($p=0.00$). Similarly, the *average time on task* (middle sub-table), as well as the *average number of queries* (lower sub-table) were lowest for task 1, and highest for task 3. Paired t-tests showed that the differences between task 1 and task 3, as well as between task 2 and task 3 were statistically significant ($p=0.05$ and $p=0.00$ respectively).

Average Perceived Complexity				
Task 1	2.74/5	t-test (p-value)		
Task 2	3.27/5	T1,T2	T2,T3	T1,T3
Task 3	3.35/5	0.00	0.74	0.00
Average Time on Task				
Task 1	0:09:52	t-test (p-value)		
Task 2	0:10:37	T1,T2	T2,T3	T1,T3
Task 3	0:12:21	0.27	0.05	0.00
Average Number of Queries				
Task 1	5.08	t-test (p-value)		
Task 2	5.34	T1,T2	T2,T3	T1,T3
Task 3	6.29	0.44	0.00	0.00

Table 5-7. Overall task characteristics

Task Assistance Results (H1)

In order to evaluate the task assistance of the different composition prototypes (noted as C1, C2 and C3), an analysis was performed on the relative task completion times, number of queries, number of page views, task success scores, as well as related questionnaire answers. T-tests were performed across systems in order to measure the

statistical significance of results, whereby $p(C1,C2)$ denotes the significance value between composition 1 and composition 2, $p(C2,C3)$ between composition 2 and composition 3 and $p(C1,C3)$ between composition 1 and composition 3.

The analysis of task times revealed that there was no statistically significant difference between the three systems, with users requiring on average 11:45 (mm:ss) per task for C1, 10:41 for C2 and 10:24 for C3 ($p(C1,C2)=0.14$, $p(C2,C3)=0.71$, $p(C1,C3)=0.09$).

However, when analysing the number of queries performed with each of the composition systems, significant differences could be noticed for C1 compared to C2 and C3. On average, users issued 6.09 queries per task with C1, 5.16 queries with C2 and 5.46 queries with C3 ($p(C1,C2)=0.00$, $p(C2,C3)=0.43$, $p(C1,C3)=0.05$) (see Figure 5-11).

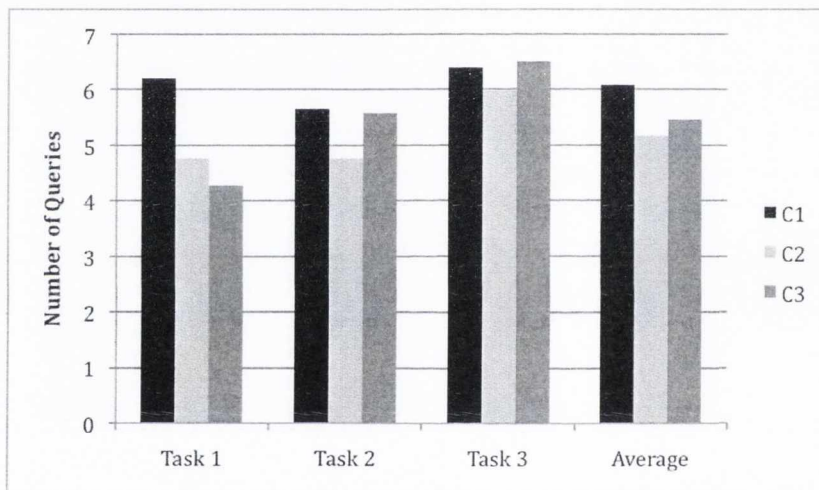


Figure 5-11. Number of Queries

This result suggests that users exhibited different search behaviours for different composition systems. More specifically, due to the less integrated structuring of results in C1, it appears that users were more likely to revert to the traditional search paradigm of frequent query reformulation. Moreover, the difference was most noticeable for tasks 1 and 2. This may be explained due to the nature of the questions in the different tasks. In particular, tasks 1 and 2 required more of a synthesis of information from different sources in order to deduce the final answers, whereas task 3 was more based around individual problem statements. The increased integration of information of C2 and C3 may therefore have provided better assistance for these tasks.

Similar results were also observed for the questionnaire question "*I found content that was relevant to my task easily.*". The average Likert score for this question was 3.23 for

C1, 3.76 for C2 and 3.38 for C3 ($p(C1,C2)=0.01$, $p(C2,C3)=0.04$, $p(C1,C3)=0.46$). These answers show that users generally felt that C2 provided the most relevant results, followed by C3 and C1.

Overall, these results therefore provide partial evidence for H1.1, revealing that users' search (i.e. query) sessions are generally observed and perceived to be more efficient with composition systems that provide increased information grouping, structuring and sequencing. However, it has to be noted that in both C2 and C3, users switched between two result screens and also used additional navigational aids (e.g. tabs). This required a slight increased user effort in terms of additional navigation clicks for these systems, namely on average 7.47 clicks for C2 and 4.70 clicks for C3.

In terms of information viewed by users, there was again a significant difference between the three composition systems. On average, users viewed 7.44 pages with C1, 9.46 with C2 and 7.96 pages with C3 ($p(C1,C2)=0.00$, $p(C2,C3)=0.00$, $p(C1,C3)=0.36$) (see Figure 5-12).

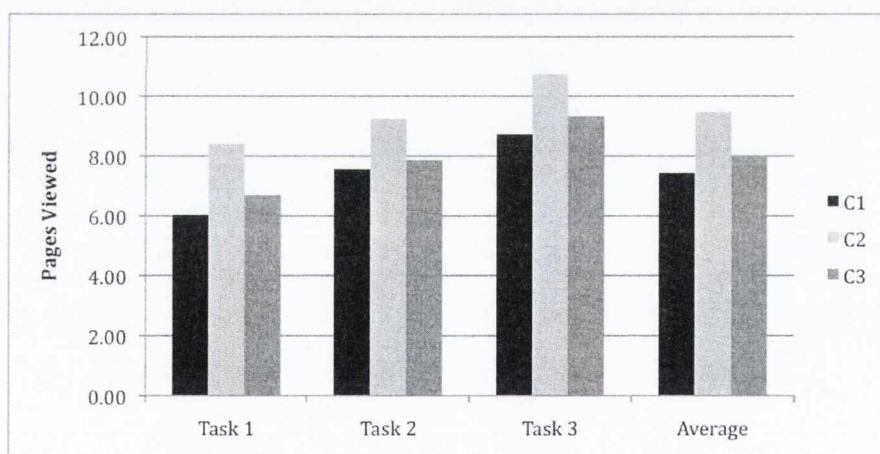


Figure 5-12. Information Viewed

This result shows that the integration of information sources and the additional navigation aids of C2 helped (and motivated) users to view an increased amount of information. Coupled with the fact that users did not end up spending more time on the task (most likely due to the decreased number of user queries), this is an encouraging result for the usage of such highly structured information presentations.

When correlating this finding with related questionnaire questions, it was revealed that users of all three systems were satisfied with the relevancy of returned information. In particular, when asked to agree or disagree with the statement *"I found the search*

system returned relevant content more often than irrelevant content.”, the average Likert score was 3.5 for C1, 3.88 for C2 and 3.80 for C3 ($p(C1,C2)=0.14$, $p(C2,C3)=0.76$, $p(C1,C3)=0.23$). Similarly, when asked to agree or disagree with the statement “I had to search a lot before I found relevant content.”, users of all systems disagreed on average, with an average score of 2.76 for C1, 2.53 for C2 and 2.74 for C3 ($p(C1,C2)=0.39$, $p(C2,C3)=0.46$, $p(C1,C3)=0.92$). These results confirm that the increased page views in C2 and C3 were not due to irrelevant information being returned, but rather due to an increased facilitation and motivation to explore additional resources.

Overall, these results therefore provide evidence for H1.2., showing that there are significant differences between the amount of information viewed for different composition systems.

In terms of task success rates, there were only slight differences between the three composition systems, which were not found to be statistically significant. On average, users had a success rate of 74.72% for C1, 79.93% for C2 and 77.74% for C3 ($p(C1,C2)=0.18$, $p(C2,C3)=0.55$, $p(C1,C3)=0.38$). However, marginally significant results were observed for the answers to the task-specific statement “The task was complex”. On average, the Likert score was 3.24 for C1, 2.88 for C2 and 3.22 for C3 ($p(C1,C2)=0.02$, $p(C2,C3)=0.05$, $p(C1,C3)=0.95$) (see Figure 5-13). A trend that can be observed in this figure is the fact that users of C2 perceived tasks 1 and 2 to be less complex, whereas the difference was less noticeable for task 3 (the most difficult task).

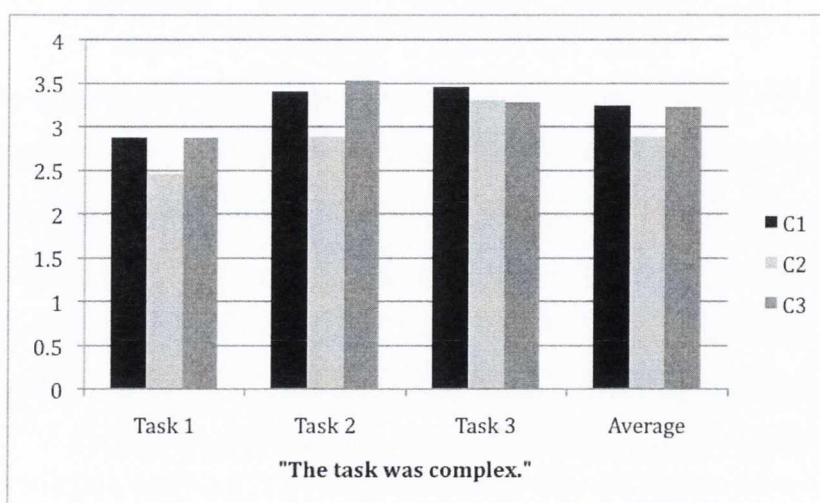


Figure 5-13. Perceived Task Complexity

Similarly, marginally significant results were observed for the answers to the statement *"I did well on the task"*. On average, the Likert score was 3.17 for C1, 3.44 for C2 and 3.19 for C3 ($p(C1,C2)=0.07$, $p(C2,C3)=0.11$, $p(C1,C3)=0.87$) (see Figure 5-14).

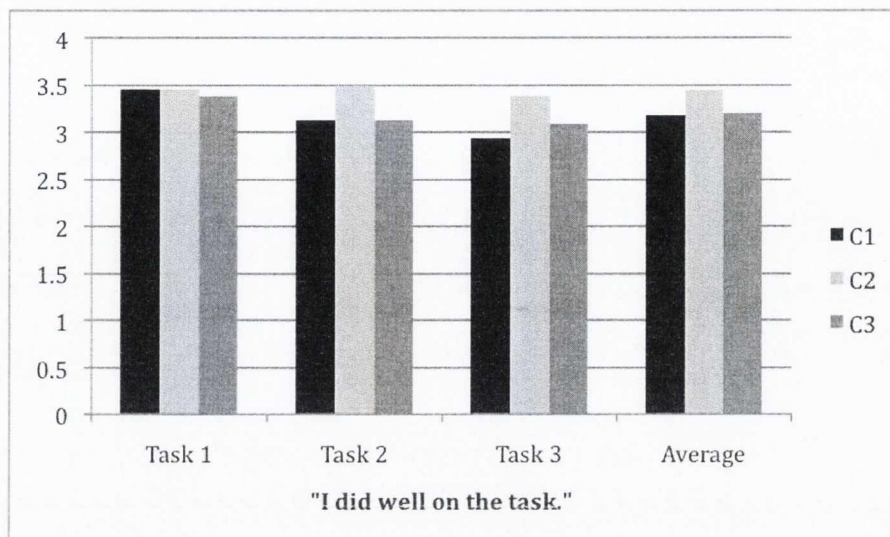


Figure 5-14. Perceived Success

A trend that can be observed in this figure is that C2 and C3 seem to have provided more consistent support across tasks 1-3, whereas the success rates of C1 dropped off more dramatically with increasing task difficulty. The above results can therefore only partially support H1.3. as only marginally significant differences were observed in terms of perceived task success rates.

When asked specifically about the system's task assistance (*"I found the result pages generated for me helpful in solving the task."*), users gave an average score of 3.5 for C1, 3.79 for C2 and 3.58 for C3 ($p(C1,C2)=0.06$, $p(C2,C3)=0.14$, $p(C1,C3)=0.58$). Again, these results only report marginally significant results, mostly between the highly integrated C2 and the loosely integrated C1. Similarly, when asked about task-specific guidance (*"The result pages guided me towards content that was relevant to the task."*), users of C2 gave significantly higher responses than users of C1 and C3 (average of 3.31 for C1, 3.78 for C2, 3.41 for C3, $p(C1,C2)=0.00$, $p(C2,C3)=0.01$, $p(C1,C3)=0.51$) (see Figure 5-15).

Overall, the above results therefore provide partial evidence for H1.4., as there are varying perceptions of task assistance for the different composition systems. However, it has to be noted that users of all three systems responded positively on average,

confirming that the open-web information compositions successfully provided task assistance in general.

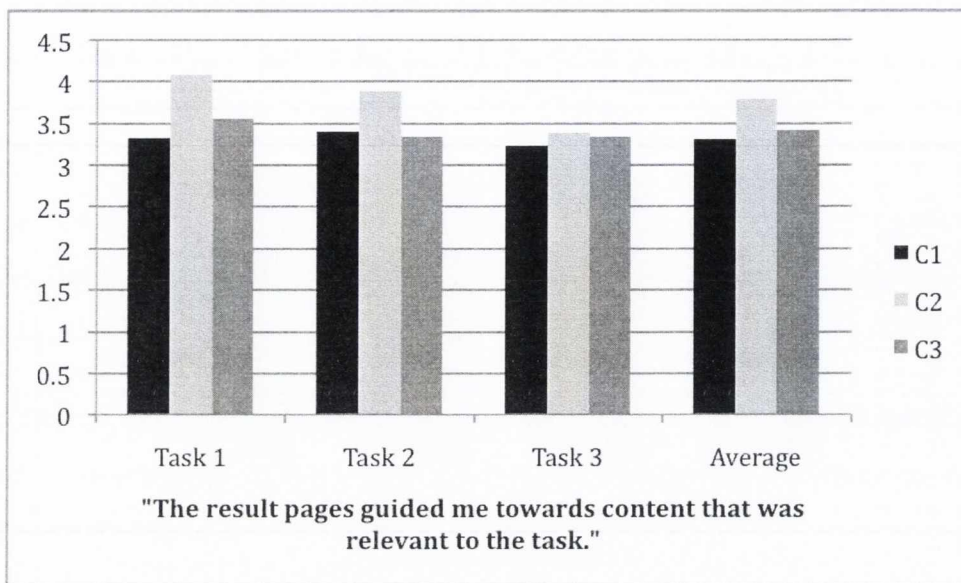


Figure 5-15. Perceived Task Assistance

User Satisfaction Results (H2)

In addition to measuring task assistance, the user study aimed at identifying users' appreciation and satisfaction regarding the various functionalities provided by the different composition systems. Moreover, user characteristics were correlated with the satisfaction scores in order to find potential interactions.

The analysis of SUS questionnaires revealed that all three systems scored similar results in terms of overall usability (73.58 for C1, 75.57 for C2, 72.95 for C3, $p(C1,C2)=0.49$, $p(C2,C3)=0.41$, $p(C1,C3)=0.89$), hence rejecting hypothesis H2.1. However, these relatively high results confirm again the general appreciation of adaptive information compositions that was already observed in the initial prototype evaluation presented in section 4.5.4.

Since no statistically significant differences could be observed between the three distinct composition systems, a number of correlation analyses were run against user characteristics that were captured in the pre-questionnaire (in order to test hypothesis H2.2). Firstly, users' overall satisfaction scores were calculated for each of the website screenshots that were shown as part of the prequestionnaire (see section 5.5.2). A Pearson correlation analysis was then run between these interface preference scores and

the measured SUS scores for the different composition prototypes. The findings were then used to determine if particular likes (or dislikes) of any of the website screenshots could be correlated to any of the three composition systems. The aim of this analysis was to measure if users' general interface preferences had any impact on their appreciation of different information composition systems (which varied significantly in terms of overall composition and presentation).

Table 5-8 presents the aggregate correlation results for each of the composition systems. As shown in this table, there are significant positive correlations between the first website screenshot and the compositions C1 and C3. These correlations can be explained by the fact that both C1 and C3 provide information compositions in a panel-based presentation (albeit to a lesser extent in C3, as information is first presented in a topic-based overview screen). Similarly, a positive correlation could be observed between C2 and the fourth website screenshot. This may again be explained by the overall similarity between the interfaces, as both C2 and the fourth screenshot provided a fully-structured result presentation that included didactical ordering.

		C1 SUS	C2 SUS	C3 SUS
WDYL wdyl.com (Panel-based information source separation)	Pearson Correlation	.442*	0.052	.379*
	Sig. (2-tailed)	0.015	0.802	0.036
	N	30	26	31
Yippy yippy.com (Ranked list, clustered navigation)	Pearson Correlation	-0.311	-0.048	0.01
	Sig. (2-tailed)	0.094	0.814	0.959
	N	30	26	31
HowStuffWorks (1) Howstuffworks.com (Ranked list, semi- structured results)	Pearson Correlation	0.135	0.01	-0.206
	Sig. (2-tailed)	0.476	0.96	0.267
	N	30	26	31
HowStuffWorks (2) Howstuffworks.com (Fully-structured result presentation)	Pearson Correlation	-0.365	.329*	-0.027
	Sig. (2-tailed)	0.087	0.047	0.889
	N	30	26	31

*. Correlation is significant at the 0.05 level (2-tailed).

Table 5-8. Interface Preference Correlations

Correlation analyses were also run against additional user characteristics that had been gathered from the pre-questionnaire. In particular, characteristics were used for

correlation analysis if a relatively equal preference split had been observed (in order to ensure sufficient numbers for each preference per composition prototype).

Table 5-9 provides an overview of the correlation results for this analysis, showing that there were generally no interactions between user characteristics and usability scores. However, one noticeable difference can be observed for composition C1, where users reported significantly higher satisfaction scores if they had indicated “*Once I understand the whole thing, I see how the parts fit.*”. This result suggests that users with global (Felder and Silverman, 1988) learning styles may have a more positive reaction towards this type of composition, whereas more sequential users seem to have a more negative view of this composition.

		C1 SUS	C2 SUS	C3 SUS
Self-reported domain expertise	Pearson Correlation	-.003	.195	.190
	Sig. (2-tailed)	.986	.340	.305
	N	30	26	31
Self-reported search Expertise	Pearson Correlation	.081	.308	.222
	Sig. (2-tailed)	.671	.126	.230
	N	30	26	31
Quick-Fix vs Understanding a problem	Pearson Correlation	-.081	.021	.094
	Sig. (2-tailed)	.672	.920	.617
	N	30	26	31
Understanding through individual parts vs Understanding through holistic view	Pearson Correlation	.455*	.160	-.067
	Sig. (2-tailed)	.013	.454	.718
	N	29	24	31

*. Correlation is significant at the 0.05 level (2-tailed).

Table 5-9. Additional characteristic correlations

Overall, the correlation analyses above have revealed that there are indeed significant differences between the three composition prototypes. In particular, it has been shown that specific interface preferences can have significant impacts on overall usability scores. These findings provide evidence for H2.2 and suggest that adaptive systems can be tailored in terms of their overall composition and presentation in order to suit particular user characteristics and preferences.

In addition to the standard usability questions presented above, questionnaires also asked users about their experience regarding adaptive navigation, composition and presentation, as well as their perceived motivation/frustration in using their respective composition system. Table 5-10 presents the overall results for the three composition prototypes in terms of navigation, composition, and presentation. This table shows that users generally found all three systems to be easy to navigate, as well as to provide clean result pages. Moreover, users did not feel overwhelmed by the amount and diversity of information and did not find the compositions unclear or inconsistent.

Q1. "The generated result pages were easy to navigate."				
C1	3.63	t-test (p-value)		
C2	3.88	C1,C2	C2,C3	C1,C3
C3	3.52	0.30	0.16	0.67
Q2. "I found the generated result pages to be clean."				
C1	3.53	t-test (p-value)		
C2	3.92	C1,C2	C2,C3	C1,C3
C3	3.29	0.17	0.02	0.37
Q3. "I found the amount and diversity of content overwhelming on every screen."				
C1	2.70	t-test (p-value)		
C2	2.77	C1,C2	C2,C3	C1,C3
C3	2.81	0.80	0.88	0.72
Q4. "I found the way the result pages were composed and presented unclear and inconsistent."				
C1	2.20	t-test (p-value)		
C2	2.35	C1,C2	C2,C3	C1,C3
C3	2.65	0.60	0.26	0.13

Table 5-10. Perceived Navigation, Composition and Presentation

Paired t-tests revealed that there was a significant difference between composition prototypes C2 and C3 for the statement *"I found the generated result pages to be clean."*, indicating that the loose topic-based structure of C3 might have had a negative impact on the overall interface usability.

In terms of query intent adaptation, users of all three systems responded positively on average to both statements *"I often used the query intent option (i.e. "I want to...") to*

narrow down the search results.” and *“The system generated appropriate presentations for the chosen query intents.”* (see Table 5-11). These results confirm that all three systems successfully performed adaptive composition and presentation to generate appropriate responses to varying user information intents. A significant difference could be found between systems for the statement *“I felt guided across the different content sources.”*, where users gave significantly higher scores for C2 compared to C1 and C3 (see Q7 in Table 5-11). This points to the fact that the increased adaptive result grouping and sequencing in C2 has positively affected users’ perceived system guidance. However, it is worth noting that users also still responded positively on average for both C1 and C3.

Q5. “I often used the query intent option (i.e. "I want to...") to narrow down the search results.”				
C1	3.33	t-test (p-value)		
C2	3.46	C1,C2	C2,C3	C1,C3
C3	3.29	0.65	0.59	0.89
Q6. “The system generated appropriate presentations for the chosen query intents.”				
C1	3.36	t-test (p-value)		
C2	3.57	C1,C2	C2,C3	C1,C3
C3	3.48	0.30	0.66	0.60
Q7. “I felt guided across the different content sources.”				
C1	3.03	t-test (p-value)		
C2	3.61	C1,C2	C2,C3	C1,C3
C3	3.22	0.00	0.05	0.29

Table 5-11. Perceived usage and appropriateness of query intent

Lastly, users generally responded positively towards the statement “Overall, I found the interaction with the system motivating” (Q7) and negatively towards “Overall, I found the interaction with the system frustrating” (Q8) (see Table 5-12). As shown in this table, there is also a significant difference between C1 and C2 for Q9, indicating that users of C2 were significantly less frustrated during the tasks. It seems therefore that the increased guidance that was observed above has had a significant impact on the perceived user experience.

However, it has to be noted again that users of all three systems responded negatively on average towards this statement for each of the composition systems, pointing

towards the fact that all three (open-web) systems were able to provide an overall positive user experience.

Q8. “Overall, I found the interaction with the system motivating.”				
C1	3.03	t-test (p-value)		
C2	3.38	C1,C2	C2,C3	C1,C3
C3	3.22	0.24	0.51	0.52
Q9. “Overall, I found the interaction with the system frustrating.”				
C1	2.60	t-test (p-value)		
C2	1.96	C1,C2	C2,C3	C1,C3
C3	2.48	0.03	0.06	0.70

Table 5-12. User motivation and frustration

Overall, the results regarding the various aspects of adaptive navigation, composition and presentation have revealed instances of significant differences between composition systems, most notably in terms of perceived guidance, interface cleanliness and user frustration (therefore providing partial evidence for H2.3).

Open-Web Information Integration (H3)

The third objective of the evaluation was to investigate the integration and composition of *open-web information*. Since the three composition prototypes made use of the open-web retrieval capabilities of the ARCHING architecture, they were each able to retrieve, compose and present such *open-web information* along with closed-corpus information.

In terms of actual open-web information usage, the analysis of search logs revealed that almost 50% of user page views consisted of open-web content (1027 out of 2143 page views). Moreover, due to the particular nature of questions in task 3 (which were more focussed on finding problem solutions), users consulted more open-web than closed-corpus resources for this task (581 out of 807 page views). This is particularly encouraging in light of the previously presented results that users on average agreed that they “*found content that was relevant (to my task) easily*” and that they “*found the search system returned relevant content more often than irrelevant content*” (see hypothesis H1). These results therefore indicate that users found relevant content across both closed-corpus and open-web information sources.

Similarly, it was shown previously that the majority of tasks were completed correctly across all three systems, indicating that users successfully used open-web resources for compiling task solutions (since many questions required information from the open-web). In addition, the fact that users on average responded positively to the statement “*I found the result pages generated for me helpful in solving the task*” and “*the result pages guided me towards content that was relevant to the task*” further confirmed that this open-web content was integrated successfully in the generated result pages. Although these results are not directly comparable to the prototype in chapter 4 (due to different task questions/difficulties), the findings nevertheless indicate that the dynamic integration of open-web content (as opposed to integrating harvested content) was similarly effective in terms of overall task assistance.

From a user satisfaction perspective, the results showed that users were generally satisfied with all three composition prototypes in terms of overall usability. In particular, the SUS analysis showed comparable results to the initial prototype presented in chapter 4, indicating that the integration of open-web information did not harm the overall system usability. Similarly, as with the prototype presented in chapter 4, users responded positively on average about their experience regarding adaptive navigation, composition and presentation, as well as the personalised guidance across different information sources. These results therefore further confirm an overall positive user experience with the various open-web prototypes, hence providing further support for hypothesis H3.

5.5.4. Overall Evaluation Findings

In summary, the evaluation results have shown that the three adaptive composition prototypes have provided (i) different degrees of task assistance (H1), (ii) different degrees of user satisfaction (H2), as well as (iii) overall positive results regarding open-web task assistance and user satisfaction.

First of all, the interaction tracking has revealed that due to the increased integration of information sources in C2 and C3, users require less query reformulations (H1.1) and are able (and motivated) to view more information (H1.2). In addition, there were also more fine-grained differences that could be observed across the three tasks. In particular, it was shown that the lower number of queries with C2 and C3 was more

significant for tasks 1 and 2 (which required more synthesis of explanatory information) than for task 3 (which consisted of individual problem statements). However, these differences in search behaviour were shown to have no impact on the overall time on task or the measured task success rate for the different composition prototypes. Nevertheless, it may be argued that composition C2 provided increased effectiveness due to the higher number of pages viewed without requiring more time. These results have also been backed up by various related questionnaire questions, showing that users perceived to find relevant information more easily in C2 compared to C1 and C3. Similarly, composition C2 was generally found to give users higher success confidence (H1.3) and that it provided improved guidance for solving the various tasks (H1.4).

In addition to task assistance, a number of satisfaction metrics have shown that users generally appreciated each of the evaluated composition systems and that they valued the respective adaptive composition, navigation and presentation functionalities. However, there were also some significant differences in terms of user satisfaction that could be found between composition systems (H2).

In particular, it was revealed that the overall usability scores of different composition types were significantly influenced by specific user characteristics and preferences (H2.2). The most notable interactions could be found between the composition types and the user preferences regarding varying website interface designs. This finding confirms that there are noticeable differences between individual user interface preferences and that such differences can have significant effects on usability scores with composition layouts. Moreover, some additional user satisfaction differences could be noticed between composition systems (H2.3) in terms of interface cleanliness and perceived frustration. These results have generally confirmed the task assistance results, namely that on average C2 was perceived to be the most satisfactory composition type.

Lastly, it is worth noting again that users of all three open-web composition prototypes responded positively on average regarding the respective systems' task assistance and usability (H3). This generally confirms the evaluation results presented in section 4.5.4, namely that the concept of adaptive information compositions successfully provides assistance for authentic user information needs. Moreover, since the prototypes presented in this chapter integrated closed-corpus as well as *open-web* information sources, it has been shown that the approach can dynamically operate over heterogeneous information sources.

In conclusion, the three composition system prototypes have each successfully applied adaptive navigation, composition and presentation techniques to support authentic user information needs. It has been shown with evidence that users of different composition types exhibit different search behaviours, and that increased adaptive composition and navigation techniques (such as in C2 and C3) lead to higher overall efficiency, effectiveness and user satisfaction. However, it has also been shown that the respective user satisfaction scores depend on particular user characteristics and preferences.

It may therefore be of benefit to the user if the most appropriate composition type was chosen based on these preferences. Since each of the composition prototypes is built on the same underlying ARCHING architecture, it is possible to flexibly switch between such composition types without requiring any system modifications. Moreover, it may be desirable to enable user-driven creations of alternative combinations of various navigation, composition and presentation techniques in order to provide even more personalised composition systems.

5.6. Conclusions

Following the successful application and evaluation of the compositional approach presented in chapter 4, this chapter has described the development of a series of distinct interface composition prototypes using the ARCHING architecture. Moreover, the results from the comparative evaluations have confirmed that the approach can successfully provide a variety of benefits to different user information needs and characteristics.

First of all, the evaluations have confirmed that users generally appreciate the adaptive composition and presentation of heterogeneous information sources and that such compositions can successfully help users in completing authentic real-life tasks. Moreover, it has been shown that users exhibit varying search behaviours for different composition types and that increased adaptation and navigation support generally enable improved user efficiency, effectiveness and satisfaction. In particular, it has been revealed that individual users' appreciation and satisfaction towards certain composition types can depend on particular user preferences and characteristics. This finding suggests that different composition types should be chosen adaptively in order to best suit an individual user's needs and preferences.

The developed prototypes have also been shown to successfully use the open-web manipulation capabilities of the ARCHING architecture, thereby achieving the dynamic integration of open-web results into adaptive information compositions. In particular, adaptive information source selection has been used successfully across the various prototypes in order to support different user intents.

In conclusion, this chapter has confirmed that the adaptive retrieval, composition and presentation of closed-corpus and open-web information can be applied successfully in order to support authentic user information needs. Moreover, it has been shown that the techniques can be applied in a number of distinct implementations in order to support particular user information needs, preferences and contexts.

6 Multilingual Information Composition & Additional Adaptation Dimensions

6.1. Introduction

The evaluation results presented in sections 4.5.4 and 5.5.3 have shown clear evidence that adaptive information compositions can successfully support users across diverse information needs, as well as heterogeneous open-web information sources. However, there are a multitude of additional adaptation dimensions that can be addressed using this compositional approach.

First of all, this chapter investigates the degree to which the compositional approach can be used to support the dimension of user language competencies. In particular, this chapter explores the notion of adaptive multilingual information compositions and the respective suitability of distinct multilingual interface compositions for different user needs and characteristics.

The second objective of this chapter is to investigate a number of additional adaptation dimensions supported by ARCHING (such as different device interfaces, user expertise modelling and multimedia preferences). Moreover, an alternative customer care application is presented, which demonstrates the transferrability of techniques to different domain and content bases without changing the underlying architectures.

The remainder of this chapter is structured as follows. Section 6.2 first presents prior work in multilingual information access systems and argues for the development of novel multilingual information presentation and interaction paradigms. Section 6.3 then presents a real-life application scenario for such multilingual information access systems, namely Personalised Multilingual Customer Care (PMCC). Section 6.4

describes a series of three multilingual composition prototypes developed using the ARCHING architecture, which aim to support bilingual users in this multilingual customer support scenario. Section 6.5 presents an evaluation of these prototypes, which reveals that bilingual users highly appreciate the composition and presentation of multilingual results. More specifically, it is shown that users particularly appreciate the composition of multilingual results in an integrated presentation, rather than providing separate presentations for different languages.

Section 6.6 then presents additional examples of adaptation dimensions that can be supported using ARCHING (including mobile device support, user expertise modelling and multimedia support), as well as the development of an alternative domain prototype. Lastly, section 6.7 concludes the chapter with a discussion of the various adaptation dimensions and prototypes, as well as the respective evaluation results.

6.2. Multilingual Information Access

As noted by Oard (2009), Multilingual Information Access (MIA) systems could be of great benefit to so-called polyglots, i.e. people who are able to at least read more than one language. A notable example of the scale of global polyglotism is the fact that “more than one billion people (i.e. 15% of the world’s population) who know at least some English are native speakers of some other language” (Oard, 2009).

Most MIA research to date has largely focussed around the notion of so-called Cross-Lingual Information Retrieval (CLIR), which typically involves the retrieval of documents in languages that are different from the original query language (Manning et al. (2008), Oard and Diekema (1998)). CLIR systems typically assume that users are unable to understand the language of the queried information sources and therefore the information delivery mainly consists of a single ranked list of translated documents/summary snippets. This fact is reinforced by the common CLIR batch-evaluation techniques that do not involve real-world users (CLEF⁵², 2000-2011), with the exception of iCLEF⁵³ (2001-2009), which mainly focuses on image search.

According to Oard (2009), major future challenges for MIA systems therefore include improvements in the overall presentation of multilingual search results, as well as the

⁵² <http://clef-campaign.org/>

⁵³ <http://nlp.uned.es/iCLEF/>

general interaction design of multilingual search systems. The opportunity for novel systems hence lies in recognising and providing such appropriate interaction support for users' multilingual competencies by moving beyond standard cross-lingual information retrieval.

One way to achieve such support is through the provision of adaptive multilingual information composition systems, which are able to dynamically compose and present multilingual responses from heterogeneous information sources. In the following sections (sections 6.3, 6.4 and 6.5), a real-world multilingual information access scenario is presented and it is demonstrated that such multilingual information compositions can indeed successfully provide adaptive support to multilingual users.

6.3. Personalised Multilingual Customer Care

As outlined in section 4.4.1, companies and organisations increasingly face challenges in addressing the various information needs of their customers, particularly given the growing diversity of user preferences and characteristics. One particularly important aspect for the success of support systems lies in the handling of multiple languages in order to serve global customers. Many organisations therefore invest significant resources in the localisation of products or services and the respective corporate support content (e.g. product documentation, online help).

However, with the growth of the social web, many customers now increasingly engage in community-driven support, for example through online forums or blogs. Such resources typically exist independently from each other, separated by their respective languages and/or regions. The most common support paradigm across these resources therefore consists of users interacting in single-language search sessions. However, the number of resources available in each language can vary substantially. For example, the English community forums⁵⁴ for the security company Symantec currently hold approximately 300,000 posts, whereas their German⁵⁵ and French⁵⁶ equivalents only hold around 15,000 and 10,000 posts respectively (note that none of these posts are translations of each other). It would therefore be highly desirable if users could avail of the various resources in a single search session.

⁵⁴ <http://community.norton.com/>

⁵⁵ <http://de.community.norton.com/>

⁵⁶ <http://fr.community.norton.com/>

6.4. Design and implementation of Multilingual Composition

Prototypes

Each of the multilingual composition prototypes presented in this section builds on the same underlying ARCHING architecture that was used for the systems described in sections 4.4 and 5.4. In particular, the prototypes are built on the same multi-model adaptation approach, which uses domain, content, user and adaptation models for the delivery of adaptive information compositions.

Since the utilised automatic metadata extraction techniques have been shown to be language-agnostic (Sah and Wade, 2011), it is possible to generate identically-structured content and domain models for the localised versions of the corporate information (i.e. the closed-corpus). Although it is currently necessary to perform manual mappings between multilingual domain concepts, significant advancements in the field of cross-lingual ontology mapping (Fu, et al., 2010) already promise to also fully automate this process in the future.

6.4.1. Multilingual Composition 1 (MC1)

The first multilingual composition prototype (MC1) follows a similar approach to the composition prototype C1 that was presented in section 5.4.1. In particular, MC1 similarly presents a “panel-based” interface, whereby information is grouped into separate panels based on the underlying information source and/or information type. However, in addition to the process presented for C1, MC1 also retrieves and composes further information based on the user’s language preferences. By translating the original query using the integrated translation capabilities (presented in section 4.3.2.3), it is possible to integrate these additional information sources in the same manner. The resulting information presentation that is displayed to the user hence consists of (i) multiple information sources and (ii) in a number of different languages.

Figure 6-1 shows a screenshot of the implemented MC1 prototype. As shown on this screenshot, both German and English results are displayed on the same screen for the German user query “aktualisieren” (i.e. updating). In this case, the user has indicated that she speaks both German and English and would therefore like to receive support information available in either of these languages. It is worth noting that English results

are only shown for the open-corpus information sources, as closed-corpus results would simply consist of the localised versions of the same content.

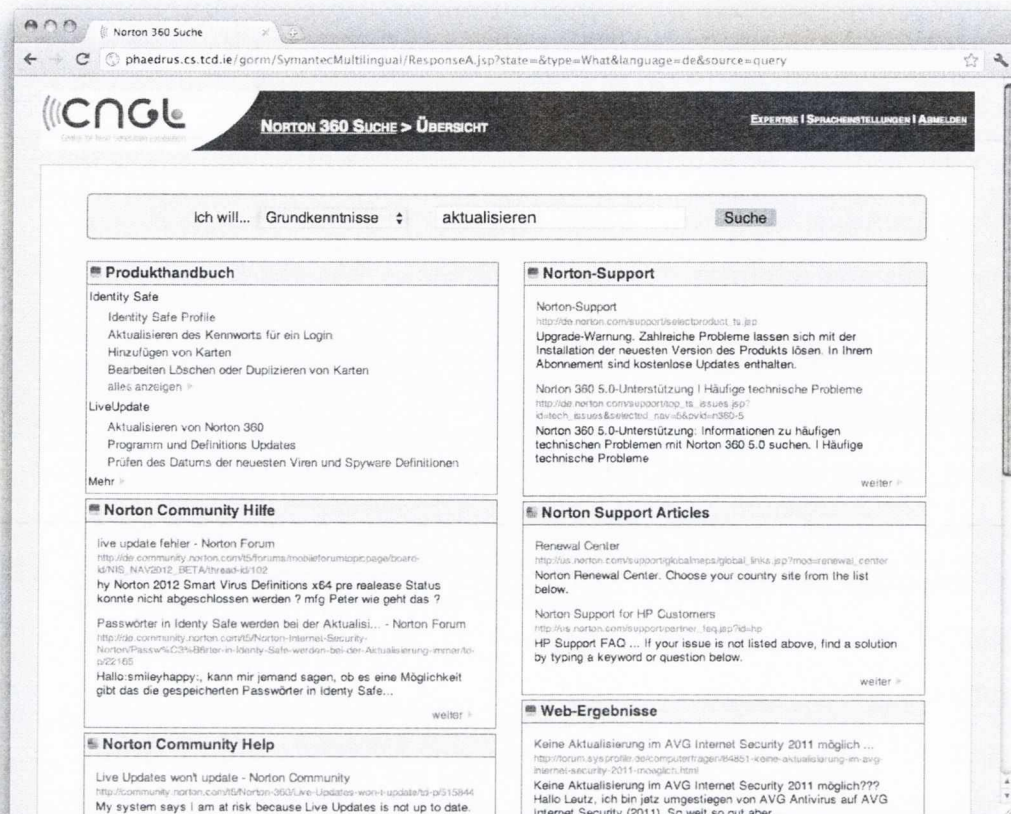


Figure 6-1. Multilingual Composition 1 (MC1) Screenshot

6.4.2. Multilingual Composition 2 (MC2)

The second multilingual composition (MC2) is based on the highly structured, topic-based composition C2 that was presented in 5.4.2. In particular, MC2 also consists of a two-stage navigation process, whereby the first stage presents an overview of initial results and groups the information according to the respective topics (i.e. product features) (see Figure 6-2). For each topic (e.g. Identity Safe), information is retrieved from the product documentation (left), as well as appropriate German and English information sources (e.g. the respective forums as shown in Figure 6-2).

After selecting one of the initial results from the first screen, the result content is displayed along with the second stage of the composition (see Figure 6-3). This second screen presents a highly structured overview of the topic related to the selected result, as well as a selection of tabs for focusing the open-corpus retrieval. As can be seen in

Figure 6-3, a user can hereby seamlessly switch between a number of German and English information sources.

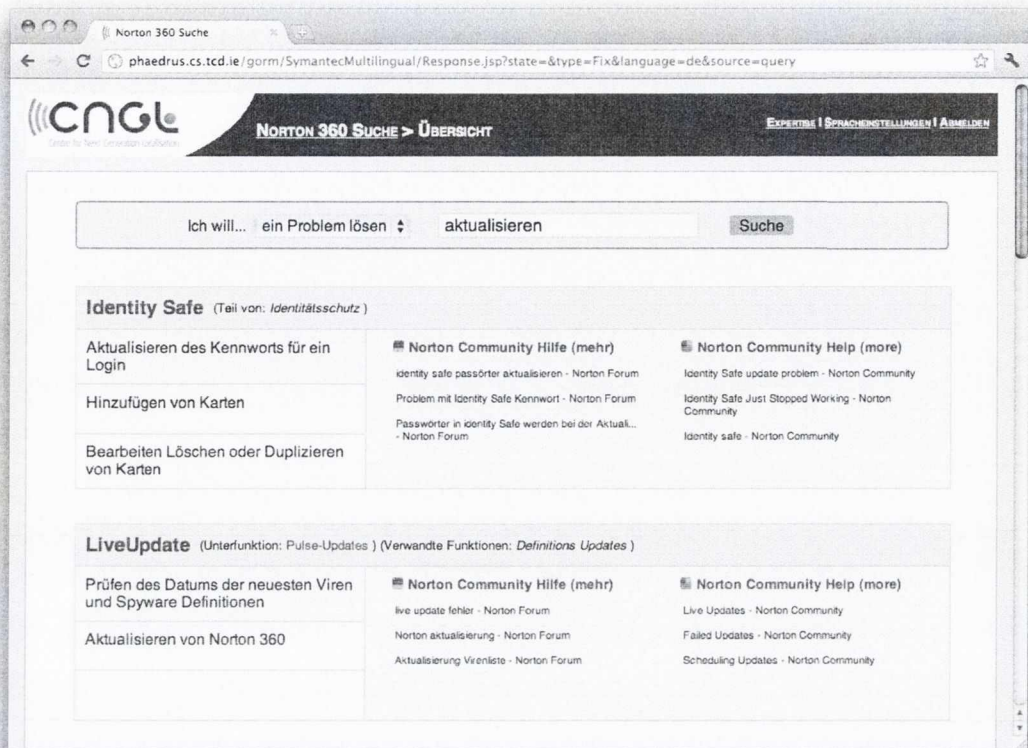


Figure 6-2. Multilingual Composition 2 (MC2) Screenshot (overview screen)

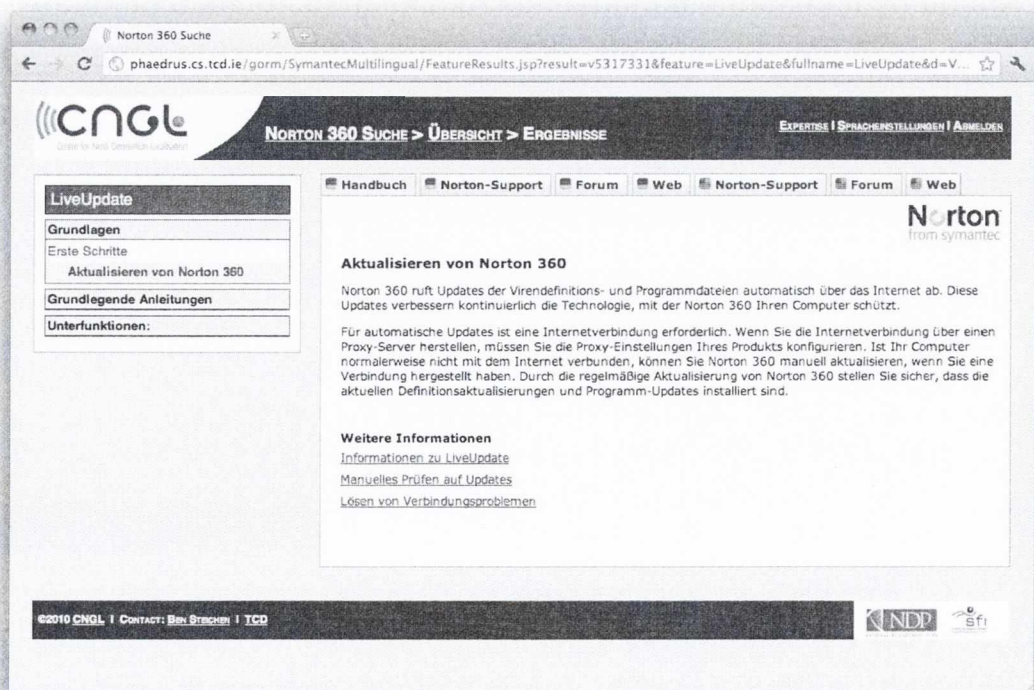


Figure 6-3. Multilingual Composition 2 (MC2) Screenshot (detailed screen)

6.4.3. Multilingual Composition 3 (MC3)

In contrast to MC1 and MC2, the third multilingual composition prototype (MC3) takes a slightly more conventional approach to multilingual information delivery. In particular, the result composition displayed by this system only presents information in a single language at any one time.

Figure 6-4 shows a screenshot of MC3, where the initial result composition only consists of results from German information sources (as this was the language of the original query, “aktualisieren”). However, as can be seen in this screenshot, users can switch between information sources by hovering over the German flag and then choosing from the list of available source languages. For example, when a user clicks on the American flag, the system completely re-generates the composition in order to present information that is exclusively sourced from English-language websites.

This type of composition thereby follows more closely the most common approach found on the web, i.e. requiring users to manually switch between pages in order to receive information in different languages.

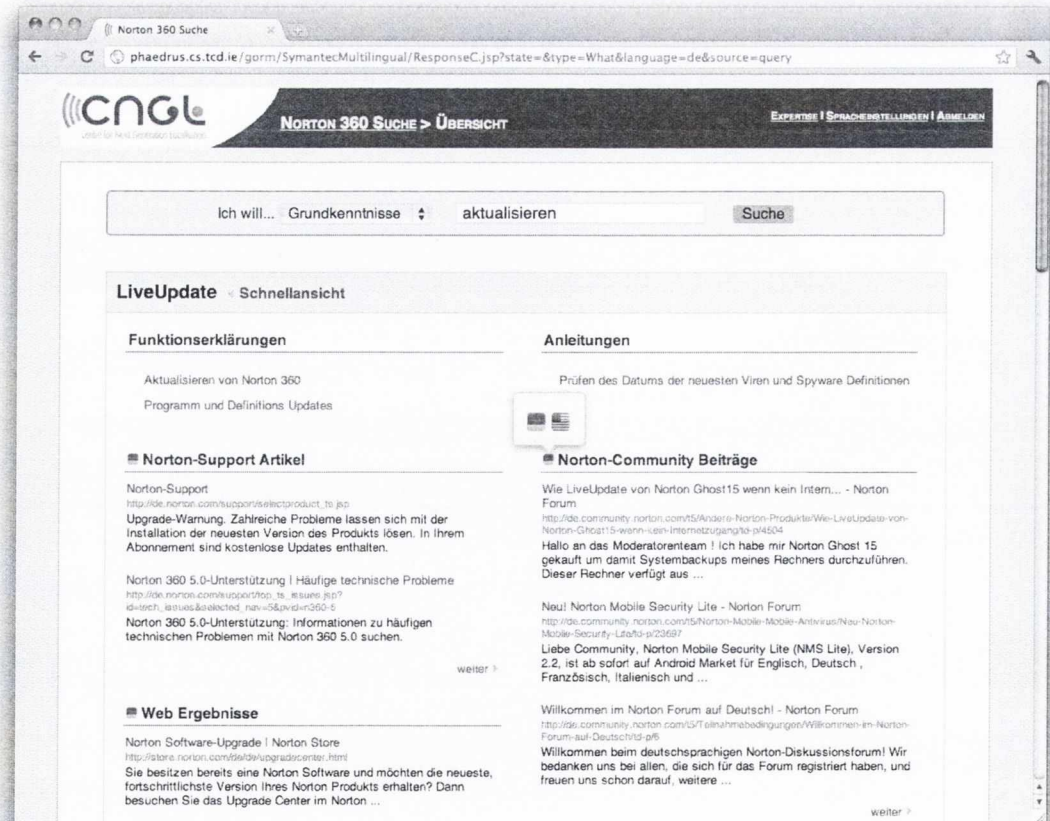


Figure 6-4. Multilingual Composition 3 (MC3) Screenshot

6.5. Evaluation of Multilingual Composition Prototypes

In order to evaluate the presented composition prototypes, a multilingual user-study was run in parallel to the (monolingual) experiment presented in section 5.5. The goal of the multilingual evaluation was again to investigate the varying degrees of task assistance and user satisfaction for different composition types. More specifically, the three prototypes were evaluated in terms of (i) their ability to support bilingual users in real-life customer support tasks (user efficiency and effectiveness) and (ii) the usability from the users' perspective (i.e. user satisfaction). The particular languages used in this study were German and English, as (i) there was significant closed- and open-corpus data available, (ii) sufficient numbers of bilingual participants could be recruited and (iii) the author of this thesis was comfortable with both languages. In order to take part in the study, participants were required to have moderate to high proficiency levels in German and English.

6.5.1. Hypotheses/Sub hypotheses

Task Assistance

Similar to the evaluation criteria presented in section 5.5.1, the benefit to the user in a multilingual customer support scenario lies in a system's ability to assist a user's search for information effectively and efficiently. In particular, a multilingual composition system should require users to invest the least amount of effort in order to find as much relevant information as quickly as possible in order to complete their task (regardless of information source language). Moreover, users should feel comfortable with the integration of multilingual resources and not be overwhelmed by the amount of information presented to them. The hypotheses regarding the user effectiveness and efficiency were therefore as follows.

- H1: The adaptive, multilingual composition types provide different degrees of task assistance.
 - H1.1: The adaptive, multilingual composition types provide different degrees of task assistance in terms of user effort for task completion.

The metrics used to validate this hypothesis are completion time and number of queries issued.

- H1.2: The adaptive, multilingual composition types provide different degrees of task assistance in terms of the amounts of relevant information viewed by users.

The metric used to validate this hypothesis is the users' overall page view count.

- H1.3: The adaptive, multilingual composition types provide different degrees of task assistance in terms of task completion effectiveness and perception.

The metric used to validate this hypothesis is the users' measured and perceived task accuracy.

- H1.4: The adaptive, multilingual composition types provide different degrees of perceived overall task assistance.

Usability questionnaire scores are used to validate this hypothesis.

User Satisfaction

In addition to the assessment of task assistance, the second goal of the evaluation was again to measure the degrees of user satisfaction for different composition types. A particular emphasis was put on the integration of multilingual information sources and the respective usefulness of the composition in a multilingual context.

The hypotheses regarding user satisfaction are therefore as follows.

- H2: The adaptive, multilingual composition types provide different degrees of user satisfaction.
 - H2.1: Overall, the multilingual composition types provide different degrees of usability scores.

Usability questionnaire scores are used to test this hypothesis.

- H2.2.: The multilingual composition types provide different degrees of usability for users with different characteristics.

In order to test this hypothesis, usability questionnaire scores are correlated with user characteristics captured during prequestionnaires.

- H2.3: Users recognise and value varying aspects of composition, adaptation and personalisation for different multilingual compositions.

In order to test this hypothesis, application-specific usability questionnaire scores are compared across compositions.

6.5.2. Experimental Setup

The experimental setup followed a similar methodology to the evaluation presented in section 5.5.2. A set of three tasks was developed for the user study, which consisted again of real-life information needs regarding the Symantec product Norton 360. Each task was set in German and contained 4 to 5 questions regarding various aspects of the product. A number of questions was designed to require users to gather information from both German and English resources in order to complete the tasks. In contrast to the (monolingual) evaluation presented in section 5.5, users in the multilingual study received each of the three composition prototypes (1 per task). This was to ensure that each system was used by significant numbers of participants, considering in particular that participant requirements were much stricter than previous studies (i.e. users were required to be comfortable in both English and German). Moreover, due to the comparative nature of this evaluation, tasks were designed to be of approximately equal difficulty in order to allow users to make fair judgements regarding system preferences. (The exact task questions, as well as the questionnaires can be found in APPENDIX G.)

As with previous evaluations, the process for each user started by receiving an e-mail about the purpose and length of the experiment, as well as the experiment URL and personal credentials. After logging in to the experiment system, users were again asked to fill out a consent form, followed by a pre-questionnaire to indicate their background regarding the Symantec product Norton 360, their general experience with search systems, as well as particular user characteristics. Moreover, users were asked about their respective language proficiencies in English and German.

Users then received instructions on how to use their first composition system (displayed as system A), including a short video tutorial and the chance to test the system using a test task. After users felt confident with the system functionalities, they would then proceed to their first task screen (displayed as task A). After the completion of their first task, users were asked to fill out a SUS questionnaire, as well as an application-

specific questionnaire which particularly focused on the integration of multilingual information sources.

Following this, users proceeded in the same manner for their second and third tasks (displayed as task B and C respectively) with their second and third systems (displayed as system B and C respectively). After completing all three tasks, users were asked to directly compare the different composition prototypes in terms of their overall usability, as well as their integration of multilingual information sources. Moreover, users were asked about general preferences in terms of multilingual information access.

As with previous evaluations, the experimental process (see Figure 6-5) was entirely online and users who had completed the entire evaluation were automatically entered into a random draw for the chance to win an electronic device. Similarly, in order to balance possible effects of order bias, tasks and systems were assigned using Latin square design.

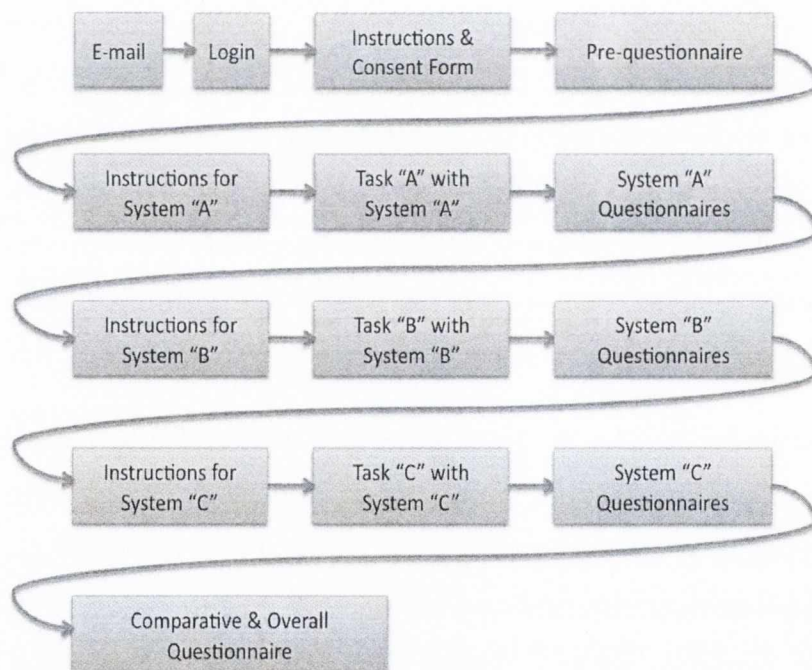


Figure 6-5. Experimental Process

6.5.3. Results

A total of 89 users signed up for the multilingual study, out of which 39 completed the full evaluation process. Participants were recruited from Trinity College Dublin, the University of South Australia, as well as Symantec Corporation.

User characteristics

As with the monolingual study, the pre-questionnaire results revealed that all users were daily computer, web and search system users. The additional user characteristics gathered from the prequestionnaire are shown in Table 6-1. As can be seen from this table, there are a number of differences among the various users.

Would you be able to advise people on using and configuring Norton 360?*	
Yes	10
No	28
Have you ever used advanced search engine features (e.g. using the '-' sign to specify unwanted terms)**	
Yes	29
No	10
What do you tend to do when encountering a software problem?	
Self-help (through manuals, forums, web searches, etc.)	33
Contact the help/call centre	5
Which of the following statements applies to you most?	
I like getting a quick how-to/fix without additional explanations.	20
I like understanding the cause of a problem that has occurred.	18
What is more important to you?	
A webpage lays out the content in clear sequential steps.	24
A webpage gives me an overall picture and relates the content to other subjects.	12
How would you generally understand new software features?	
Once I understand all the parts, I understand the whole thing.	18
Once I understand the whole thing, I see how the parts fit.	17
What is your level of proficiency in English?	
Moderate/High	24
Native	15
What is your level of proficiency in German?	
Moderate/High	12
Native	27

*. this question was aimed at estimating a user's domain expertise

** . this question was aimed at estimating a user's search expertise

Table 6-1. Overall prequestionnaire answers for the multilingual study

This information will be used below to determine user satisfaction correlations between specific user characteristics and the different composition systems.

Task Assistance Results (H1)

As with previous evaluations, in order to evaluate the task assistance of the different multilingual composition prototypes (noted as MC1, MC2 and MC3), an analysis was performed on the relative task completion times, number of queries, number of page views, task success scores, as well as related questionnaire answers. T-tests were performed across systems in order to measure the statistical significance of results, whereby $p(\text{MC1,MC2})$ denotes the significance value between composition 1 and composition 2, $p(\text{MC2,MC3})$ between composition 2 and composition 3 and $p(\text{MC1,MC3})$ between composition 1 and composition 3.

First of all, the analysis of task times revealed that there was a significant difference between MC3 compared to MC1 and MC2, with users requiring on average 14:48 (mm:ss) per task for MC1, 14:20 for MC2 and 16:50 for MC3 ($p(\text{MC1,MC2})=0.40$, $p(\text{MC2,MC3})=0.02$, $p(\text{MC1,MC3})=0.05$). This finding may be explained by the fact that users of MC3 had to use the “language switch” feature in order to view alternative language resources. This (more conventional) way of composing therefore seems to have a negative impact on user efficiency in terms of task completion times. As can be seen in Figure 6-6, this result could be observed across each of the three tasks.

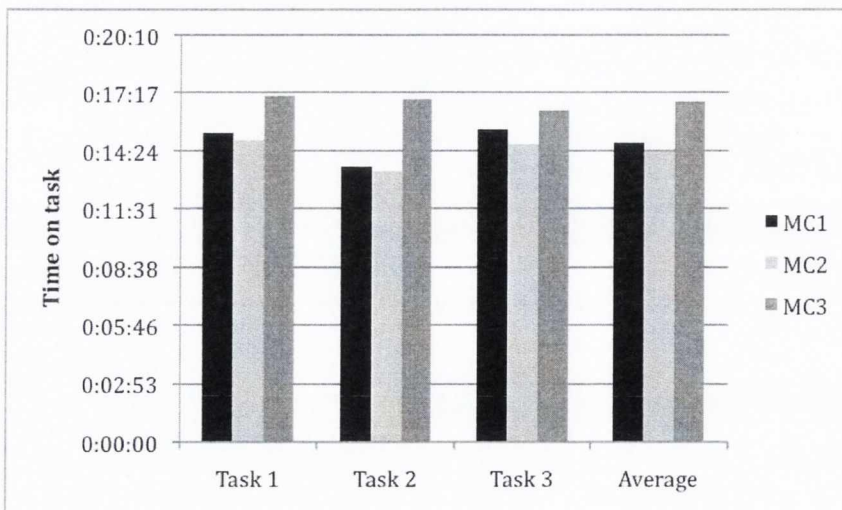


Figure 6-6. Time on task

Similar to the evaluation results observed for the monolingual study (section 5.5.3), multilingual composition users also exhibited varying search behaviours in terms of

number of queries. In particular, when using MC2, users performed significantly fewer queries than with MC1 or MC3. On average, the number of queries was 9.93 per task for MC1, 8.31 for MC2 and 10.25 for MC3 ($p(\text{MC1,MC2})=0.02$, $p(\text{MC2,MC3})=0.01$, $p(\text{MC1,MC3})=0.82$) (see Figure 6-7).

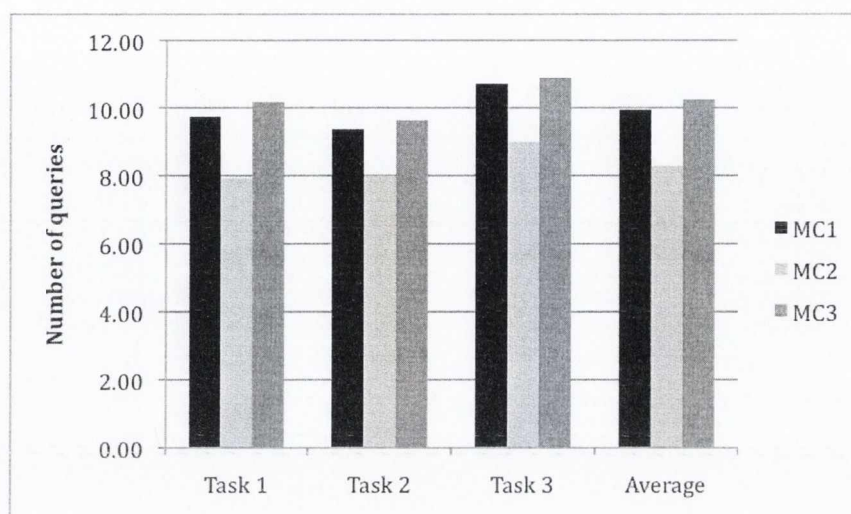


Figure 6-7. Number of queries

This points again to the fact that the increased integration of information sources in MC2 seems to have encouraged users to browse the presented result compositions. By contrast, when using MC1 and MC3, users seem to have opted to revert to more frequent query reformulations in order to find the desired information. It is however worth pointing out that this behaviour for MC3 is different from the monolingual user study (where C2 and C3 had similar number of queries). The different composition of multilingual results therefore seems to have also had an impact on users' efficiency in terms of required number of queries. Again these findings could be observed across each of the three tasks (see Figure 6-7).

Both of these results hence provide evidence for H1.1, as the three composition types clearly provide different degrees of task assistance in terms of the time spent on tasks and the number of queries performed. Moreover, it has been shown that the multilingual aspect has had a negative effect on composition MC3, most likely due to its more conventional approach for multilingual result presentation (i.e. requiring users to switch between languages).

In terms of the amount of information viewed with each composition type, similar results to the monolingual study could be observed. On average, the number of page

views was 10.40 per task for MC1, 13.52 for MC2 and 11.76 for MC3 ($p(\text{MC1,MC2})=0.00$, $p(\text{MC2,MC3})=0.10$, $p(\text{MC1,MC3})=0.21$) (see Figure 6-8).

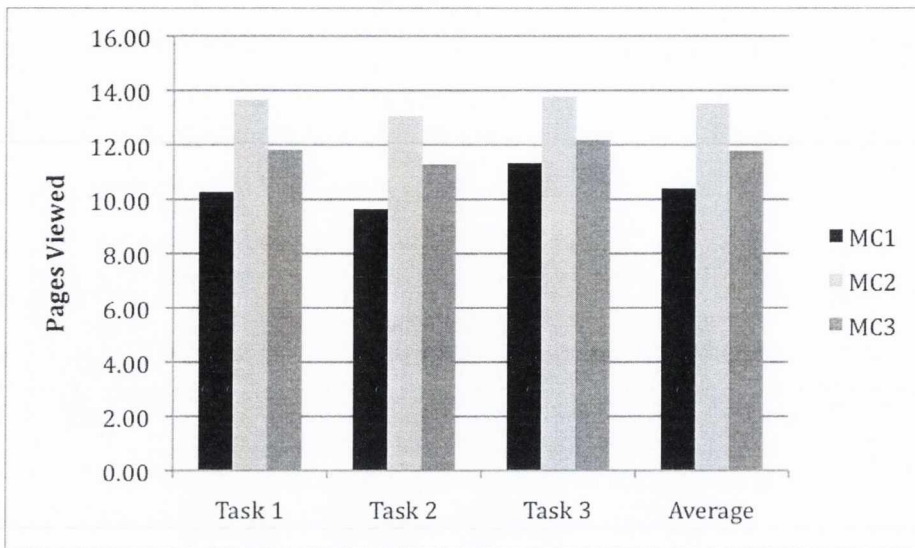


Figure 6-8. Information Viewed

This results confirms again that the increased grouping and integration of sources provided by MC2 and MC3 have had a positive effect on users' motivation to view more information. While this can be regarded as a positive finding for MC2 (as there was no increase in time spent), it may be argued that MC3 may have been less effective overall (since users spent significantly more time overall).

These findings overall provide evidence for hypothesis H1.2, showing that the multilingual composition types provide different degrees of task assistance in terms of the amounts of relevant information viewed by users.

In contrast to the findings from the monolingual evaluation, the results in terms of task success provided significant differences for the different multilingual composition types, in particular between MC3 compared to MC1 and MC2. On average, users had a success rate of 74.16% for MC1, 78.47% for MC2 and only 64.09% for MC3 ($p(\text{MC1,MC2})=0.26$, $p(\text{MC2,MC3})=0.00$, $p(\text{MC1,MC3})=0.05$) (see Figure 6-9).

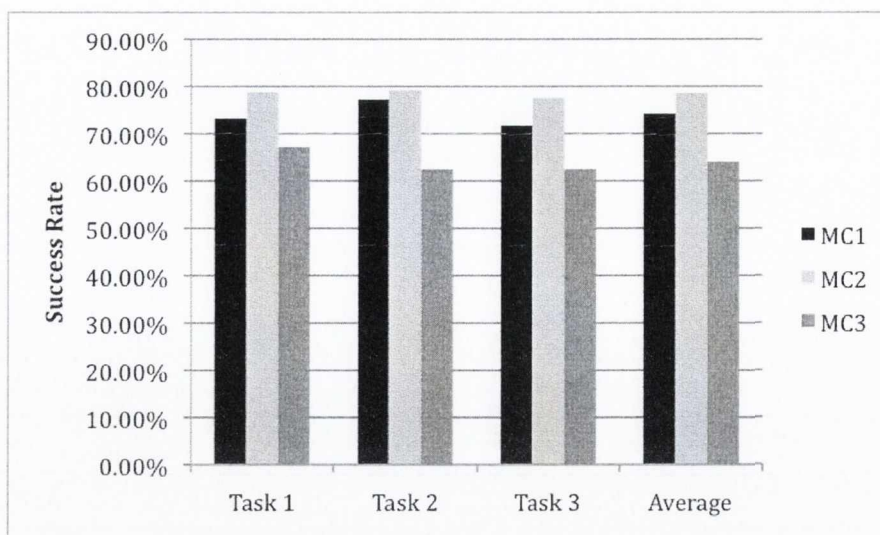


Figure 6-9. Task Success

This reinforces the fact that the approach of providing separate information presentations for different languages can have an overall negative impact on efficiency, as well as effectiveness (H1.3).

Lastly, these findings were also confirmed when asking users about their perceived task assistance for the respective composition systems. As shown in Table 6-2 and Figure 6-10, when using MC3, users were less confident of their success and generally found the composition to be less helpful in solving the tasks. In particular, significant differences could be found for the question “I did well on the task” and “The result pages guided me towards content that was relevant to the task”. Moreover, a general trend could be observed that users also favoured composition MC2 over MC1.

These results therefore provide evidence for hypothesis H1.4, since the three multilingual composition types provide different degrees of perceived task assistance. However, it has to be noted that on average users responded positively to the various questionnaire questions (in particular for MC1 and MC2), confirming that the compositional approach successfully provided task assistance in a multilingual scenario.

Q1. "The task was complex."				
MC1	2.74	t-test (p-value)		
MC2	2.52	MC1,MC2	MC2,MC3	MC1,MC3
MC3	2.97	0.43	0.09	0.37
Q2. "I did well on the task."				
MC1	3.30	t-test (p-value)		
MC2	3.42	MC1,MC2	MC2,MC3	MC1,MC3
MC3	2.97	0.55	0.00	0.05
Q3. "I found the result pages generated for me helpful in solving the task."				
MC1	3.46	t-test (p-value)		
MC2	3.47	MC1,MC2	MC2,MC3	MC1,MC3
MC3	3.24	0.94	0.25	0.31
Q4. "The result pages guided me towards content that was relevant to the task."				
MC1	3.51	t-test (p-value)		
MC2	3.71	MC1,MC2	MC2,MC3	MC1,MC3
MC3	3.29	0.29	0.03	0.27

(Note: all values are averages across the three tasks)

Table 6-2. Perceived Task Assistance

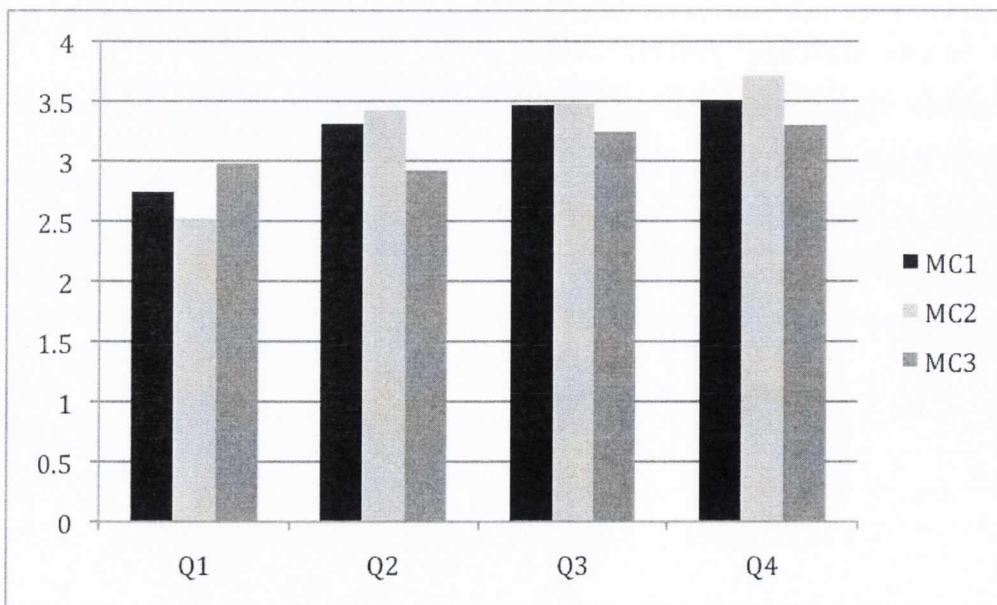


Figure 6-10. Perceived Task Assistance

User Satisfaction Results (H2)

As with previous evaluations, the multilingual user study also aimed at identifying users' appreciation and (comparative) satisfaction regarding the different composition systems. Moreover, various user characteristics were correlated with the satisfaction scores in order to find potential interactions.

First of all, the respective SUS scores were 70.59 for MC1, 71.47 for MC2 and only 64.55 for MC3 ($p(\text{MC1,MC2})=0.83$, $p(\text{MC2,MC3})=0.08$, $p(\text{MC1,MC3})=0.16$). Although these differences were not found to be statistically significant at the rigorous significance cut-off point of 0.05, they generally followed the trend that users were least satisfied with the overall usability of MC3. Similar results was also observed when asking users the question "Which system did you prefer the most?". In total, MC1 and MC2 were preferred by 14 users each, whereas MC3 was only preferred by 8 users. Similarly, when asked "Which system did you prefer the least?", there was a total of 10 users each for MC1 and MC2, whereas MC3 was preferred the least by a total of 16 users. These findings therefore generally confirm hypothesis H1.1, since the three compositions provided different degrees of overall usability scores.

In addition to finding overall user satisfaction scores, the study also aimed at identifying if certain user characteristics may have led to individual differences between systems. To this end, system preferences were aggregated for each individual user characteristic. A number of notable differences could be observed through this correlation, as shown in Figure 6-11, Figure 6-12 and Figure 6-13. First of all, users who indicated in the pre-questionnaire that they "liked getting a quick how-to/fix" strongly preferred MC1, whereas users who "like understanding the cause of a problem" strongly preferred MC2 (see Figure 6-11). This result suggests that the additional adaptive grouping, sequencing and navigation of information in MC2 has provided more comprehensive information compositions (i.e. overall more informative). A second difference that could be observed across user characteristics regarding "sequential" vs "global" learning (Felder and Silverman, 1988). Users who indicated that they like a website to "lay out the content in clear sequential steps" strongly preferred MC2 (see Figure 6-12). This finding indicates that the additional adaptive functionalities of MC2 have generally provided a more sequential information presentation overall. Similarly, differences were also observed for the related question about understanding through individual parts or more holistic learning. As already

shown in the monolingual evaluation (see section 5.5.3), users who like understanding through individual parts strongly preferred MC2, whereas people who like more holistic learning strongly preferred MC1.

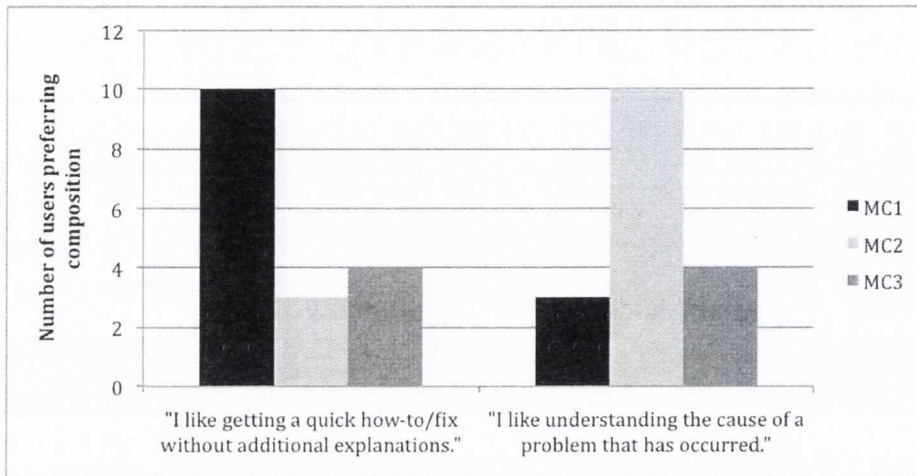


Figure 6-11. "Quick-Fix" vs "Understanding a problem"

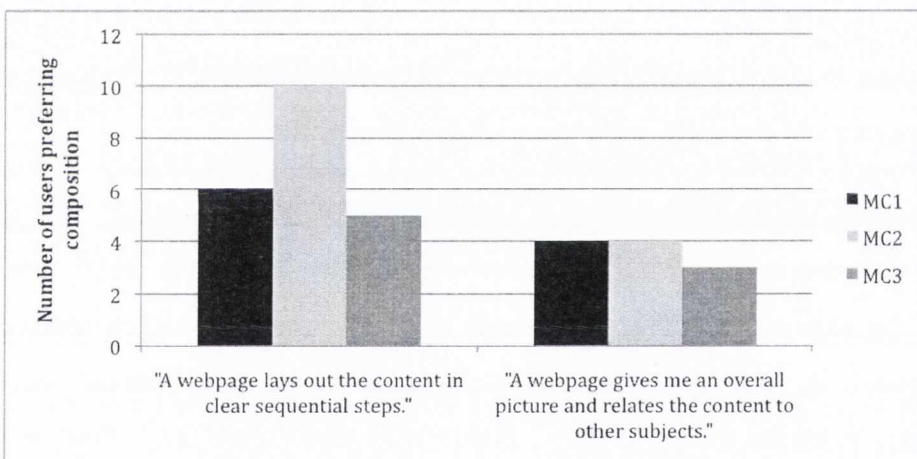


Figure 6-12. "Sequential" vs. "global" learning

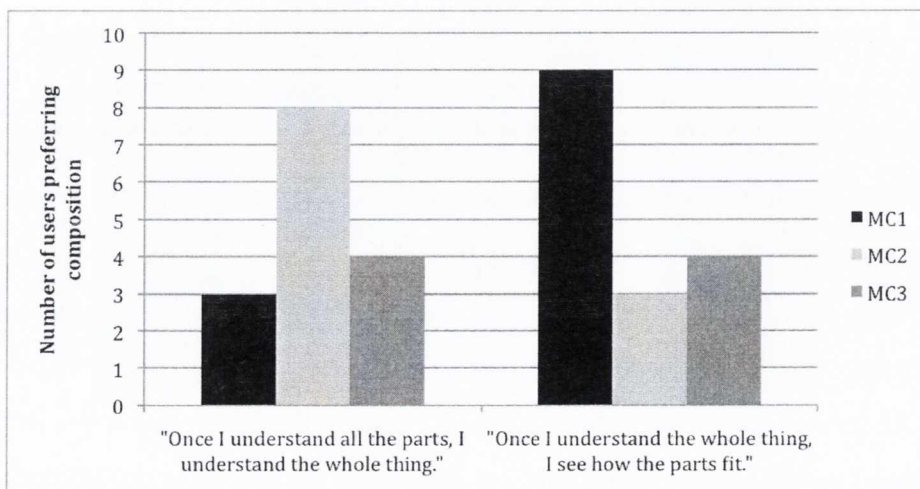


Figure 6-13. Understanding through individual parts vs holistic view

These findings provide clear evidence for hypothesis H.2.2, as they have shown that multilingual composition types provide different degrees of usability for users with different characteristics.

A number of additional, more specific questionnaire questions asked about (i) users' appreciation of the individual composition functionalities (after they had just completed a task with a prototype) and (ii) their overall composition preferences in terms of navigation, presentation and multilingual information integration (after users had used all of the composition systems). Table 6-3 and Figure 6-14 present the overall results regarding users' responses for the different composition prototypes.

Q5. "I found it easy to navigate across German and English content."				
MC1	3.76	t-test (p-value)		
MC2	3.81	MC1,MC2	MC2,MC3	MC1,MC3
MC3	3.29	0.82	0.04	0.05
Q6. "I would like the presentation of German and English content more clearly separated from each other."				
MC1	2.33	t-test (p-value)		
MC2	2.10	MC1,MC2	MC2,MC3	MC1,MC3
MC3	2.00	0.39	0.66	0.23
Q7. "I would like the German and English content to be more integrated."				
MC1	2.79	t-test (p-value)		
MC2	2.86	MC1,MC2	MC2,MC3	MC1,MC3
MC3	3.45	0.74	0.00	0.00
Q8. "I felt guided across the German and English content."				
MC1	3.25	t-test (p-value)		
MC2	3.52	MC1,MC2	MC2,MC3	MC1,MC3
MC3	2.64	0.19	0.00	0.00
Q9. "I found the German and English content to be highly related to each other."				
MC1	3.41	t-test (p-value)		
MC2	3.39	MC1,MC2	MC2,MC3	MC1,MC3
MC3	3.05	0.93	0.08	0.07

Table 6-3. User Satisfaction regarding multilingual information composition

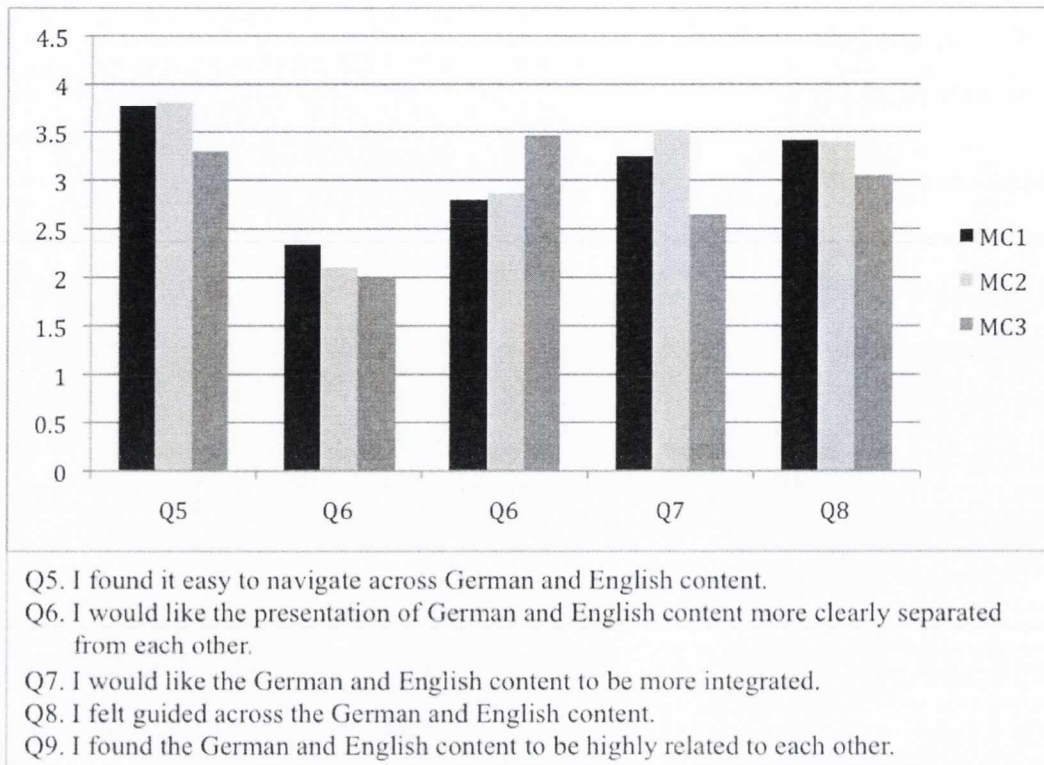


Figure 6-14. User Satisfaction regarding multilingual information composition

As can be seen from these results, users generally found it easier to navigate across German and English content with MC1 and MC2 compared to MC3. Users indicated that they would like German and English content to be more integrated in MC3. Moreover, results showed that overall MC2 provided the most guidance across German and English content. Similar results were also observed when users were asked directly about the comparative ease of navigation and overall guidance (see Figure 6-15), thereby confirming hypothesis H2.3.

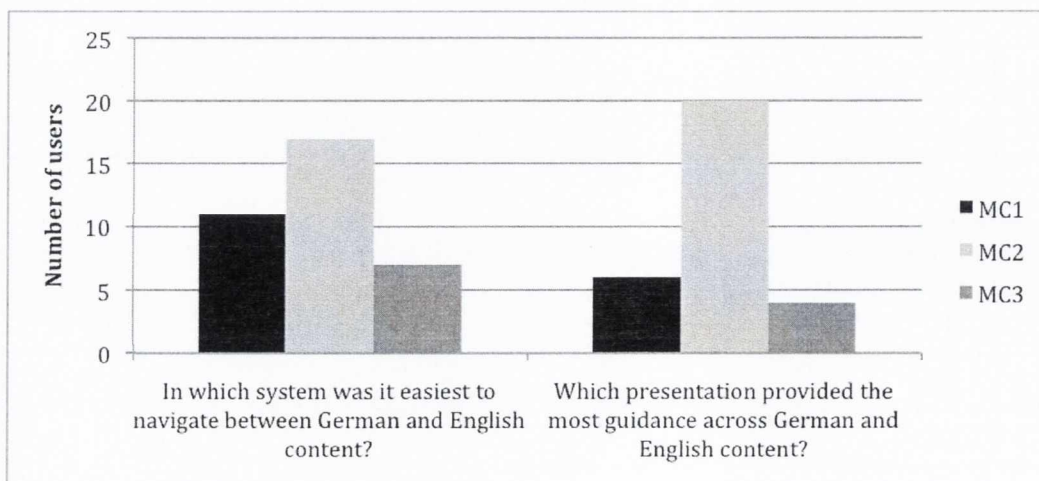


Figure 6-15. Direct comparison between multilingual compositions

Lastly, users were asked about their general multilingual information access preferences (see Figure 6-16), revealing that the majority of users are comfortable with multiple languages displayed on the same screen (Q1), that they prefer it when content is presented in its original language (Q2) and that they generally do not prefer to have all content presented in a single language (Q3).

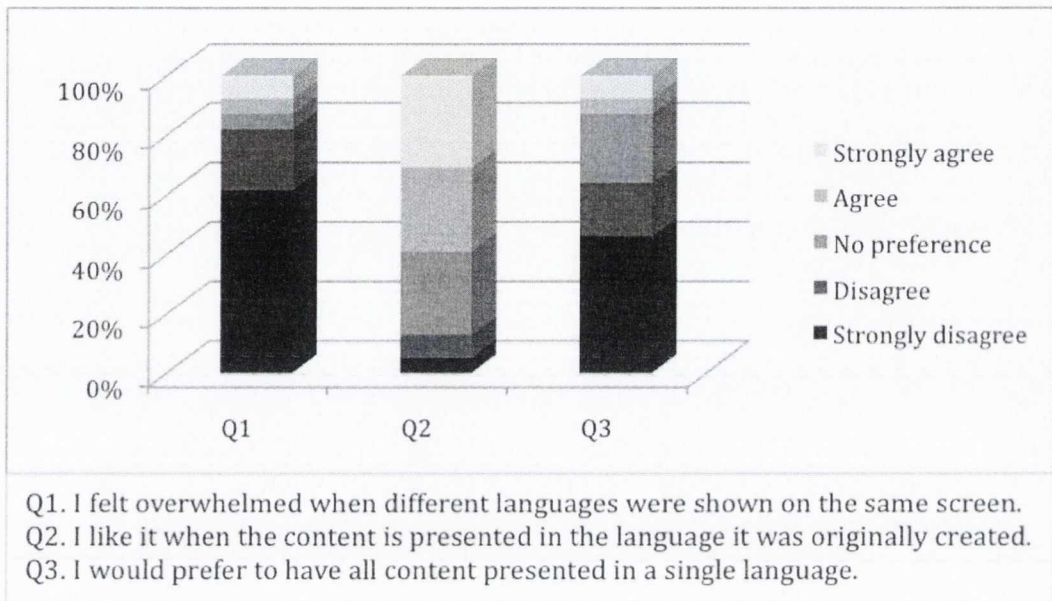


Figure 6-16. General Multilingual Information Access Preferences

6.5.4. Overall Evaluation Findings & Additional Support Possibilities

In summary, the evaluation results have shown that the three adaptive multilingual composition prototypes have provided (i) different degrees of task assistance (in terms of measured and perceived efficiency and effectiveness) (H1) and (ii) different degrees of user satisfaction (H2).

First of all, the interaction tracking has revealed that due to the more conventional multilingual information access paradigm of MC3 (i.e. using a “language switch”), users required more time (H1.1) and were less successful in completing their tasks than with the more integrated MC1 and MC2 (H1.3). Moreover, results showed that users exhibited different query behaviours, with users of MC2 (i.e. the most comprehensive and integrated composition) requiring the least number of queries (thereby confirming the results from the monolingual study in section 5.5.3) (H1.1). Moreover, due to the increased integration of information sources in MC2, users were able (and motivated) to view more information overall (H1.2). These results have also been backed up by

various related questionnaire questions, showing that users perceived MC2 (and to a lesser extent MC3) to provide improved guidance compared to MC3. Similarly, compositions MC1 and MC2 were generally found to give users higher success confidence scores (H1.4).

Overall, it is worth noting again that users of all three composition prototypes responded positively on average regarding the respective systems' task assistance (although to a lesser extent for MC3). This generally confirms that the concept of adaptive information compositions can successfully provide assistance for authentic, multilingual user information needs.

In addition to task assistance, a number of satisfaction metrics have shown that users overall appreciated each of the evaluated composition systems and that they valued the respective adaptive composition, navigation and presentation functionalities. However, there were also some significant differences in terms of user satisfaction that could be found between composition systems (H2).

Similar to the task assistance findings, the results regarding user satisfaction showed that users overall preferred MC1 and MC2 compared to MC3. Moreover, it was revealed that the overall usability scores of different composition types (and the corresponding system preferences) were significantly influenced by specific user characteristics and preferences (H2.2). The most notable interactions could be found between the composition types and the user's general approach to problem solving (i.e. "Quick-Fix" vs "Understanding a problem"), as well as users' learning styles (i.e. "Sequential" vs. "global" learning and "Understanding through individual parts" vs "Understanding through holistic view"). Moreover, some additional user satisfaction differences could be noticed between composition systems (H2.3) in terms of ease of multilingual composition navigation, integration and perceived guidance. These results have generally confirmed the task assistance results, namely that on average MC2 was perceived to be the most satisfactory composition type, and that MC3 was perceived to be the least satisfactory. Again, this is most likely due to the separation of languages in this composition types, as users have shown strong support for the tight integration of results, irrespective of the information source language.

In conclusion, the three composition system prototypes have each successfully applied adaptive navigation, composition and presentation techniques to support authentic multilingual user information needs. It has been shown with evidence that users of

different composition types exhibit different search behaviours, and that the increased integration of multilingual information of MC1 and MC2 have lead to higher overall efficiency, effectiveness and user satisfaction.

However, it has also been shown that the respective user satisfaction scores between the two most satisfactory composition types (MC1 and MC2) depend on particular user characteristics and preferences. As with the monolingual composition systems, it may therefore be of benefit to the user if the most appropriate composition type was chosen based on these preferences. Again, since each of the composition prototypes is built on the same underlying ARCHING architecture, it is possible to flexibly switch between such composition types without requiring any system modifications.

While this study has revealed that the application of multilingual information compositions can indeed be successful for bilingual users, there are many additional opportunities for applying such techniques for users with different levels of language proficiency. In particular, as users may have lower capabilities in certain languages (e.g. beginner or moderate), different types of compositions with varying adaptive navigation techniques could be of particular benefit. For example, users with moderate skills in a particular language might be comfortable with identifying relevant content in that language (e.g. from user forums), but may require additional support in order to fully comprehend the complete topic. In this case, an on-the-fly translation option could be of great benefit to the user, allowing the study of the original information, as well as its automated translation. Figure 6-17 shows an implementation of such a prototype, where a French-speaking user is currently reading an English forum thread. In this screenshot, the user has previously chosen to translate the first post into French and is currently requesting for an additional on-the-fly translation of the third post.

Moreover, users with multiple language skills (i.e. users who understand more than 2 languages) could be supported by adaptively retrieving, composing and presenting information from each of the user's preferred language sources. Figure 6-18 shows an example of such a multilingual composition, where the user has previously indicated language proficiencies in French (preferred), English and German.

This multilingual information composition approach may also be of benefit in a number of different application scenarios (e.g. news domains), enabling balanced compositions of diverse viewpoints from multiple languages and cultures.

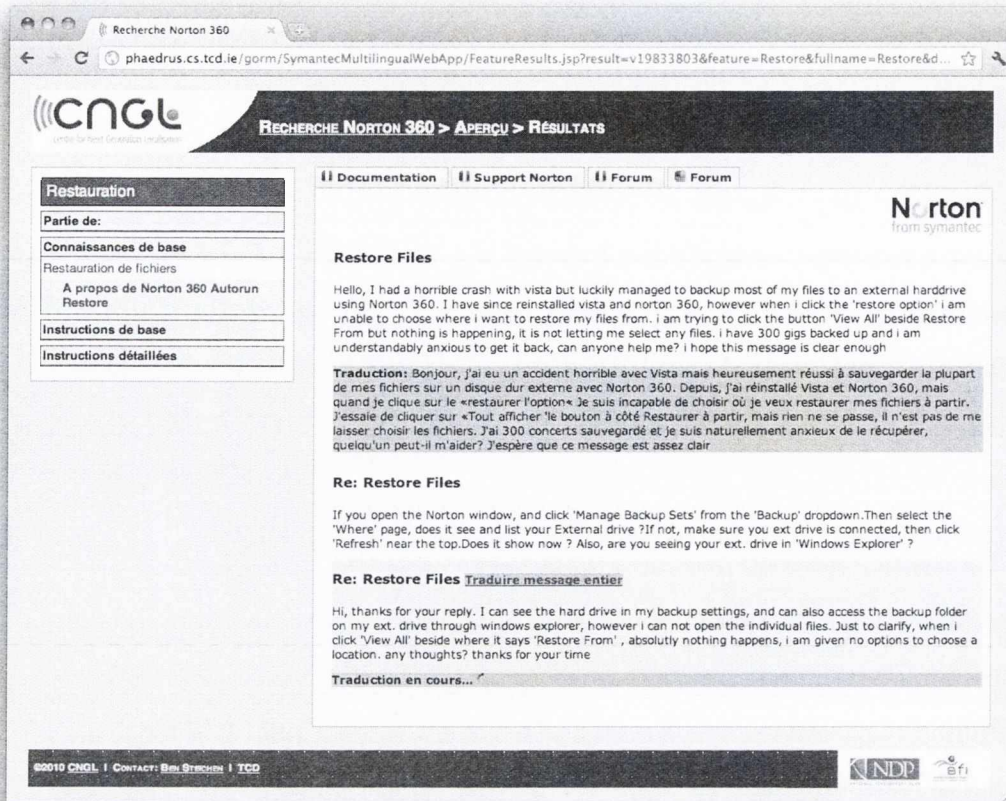


Figure 6-17. Multilingual Composition with translation support

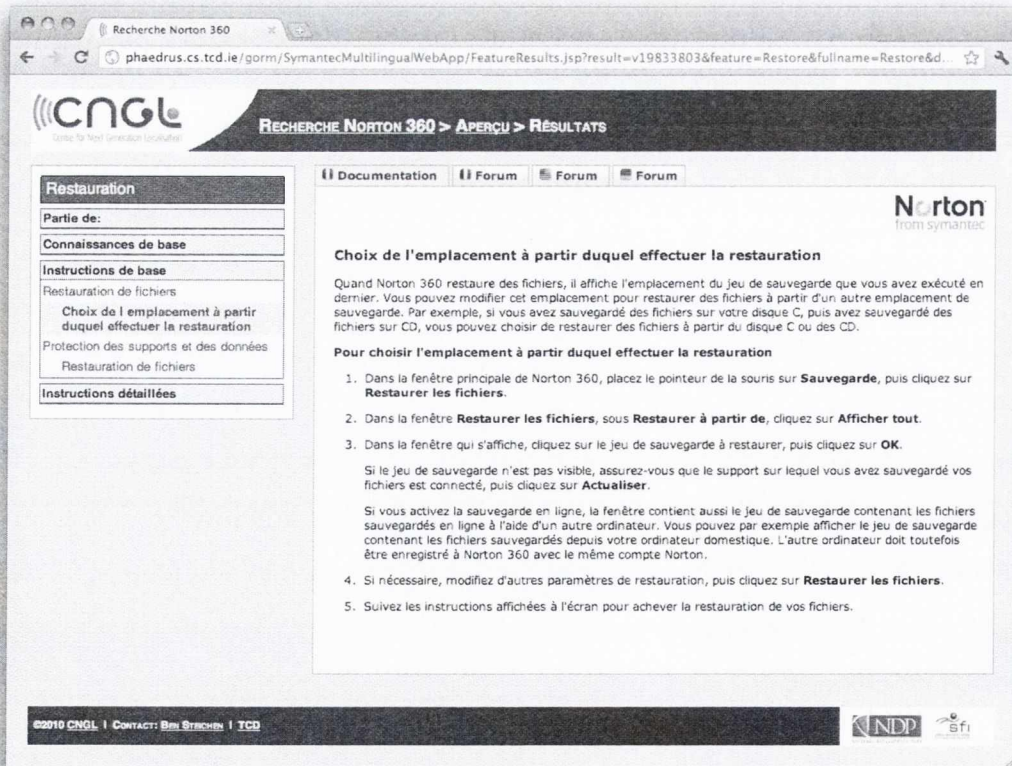


Figure 6-18. Multilingual Composition with French, English and German results

6.6. Additional Adaptation Dimensions and Prototypes

In addition to the multilingual capabilities presented in the prototypes and evaluations above, the compositional approach can be applied for a multitude of further adaptation dimensions and applications. This section presents some of the additional implementations that have been developed using the presented techniques and technologies, showing a range of possibilities for further work in this area.

In particular, section 6.6.1 presents a prototype, which uses the adaptive composition and presentation techniques to adapt to mobile devices. Section 6.6.2 presents an implementation of more fine-grained user expertise modelling, including an interface for users to manually review and change such models. Section 6.6.3 shows how the open-corpus retrieval capabilities can be used to include additional media types (e.g. video), in order to support varying user preferences. In section 6.6.4, an alternative customer care application is presented, which demonstrates the transferrability of techniques to different domain and content bases without the need for changes to the underlying architectures.

6.6.1. Mobile device adaptation

As presented throughout the various composition prototype descriptions above, the underlying ARCHING architecture can be used to tailor both the composition and presentation to particular user needs, characteristics and preferences. These capabilities can also be used to tailor specific compositions and presentations towards particular user device characteristics, including for example mobile devices. More specifically, compositions and presentations can be adapted to best fit different screen sizes and access modalities (e.g. touch interfaces).

Figure 6-19 presents an implementation of such a tailored composition, which is generated using (i) simplified composition strategies (i.e. composing more focused information per screen) and (ii) optimised presentation strategies (i.e. using transformation models that transform the composed result models into presentations suitable for different screens and access modalities). Figure 6-19 (a) presents a result overview screen for the query “update”, showing a simplified composition of the overall documentation topics. Figure 6-19 (b) presents a selected documentation item

(“About Program and Definition Updates”), as well as a simplified navigation to access the various other (multilingual) information sources (e.g. “Forums”).

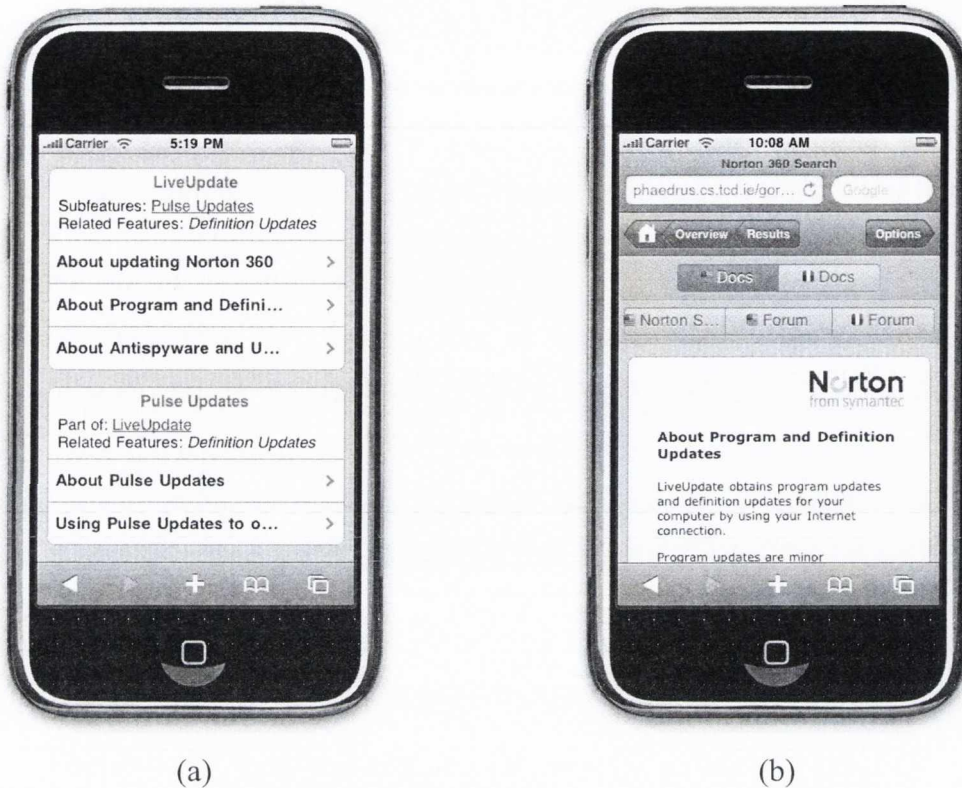


Figure 6-19. Mobile device compositions

6.6.2. Fine-grained user expertise modelling

In addition to the user variables presented in the various composition prototypes above, more fine-grained user models may also be used to provide even more personalised compositions.

For example, as users interact with composition systems, their knowledge levels regarding particular topics (e.g. product features) are constantly growing. This information can be captured by the system in order to provide more fine-grained information compositions according to the respective expertise levels. More specifically, as users browse information in the composition prototypes (either from closed- or open-corpus), the system can identify the most related product feature for this information (either through direct metadata lookup or open-corpus classification) and increase the user’s expertise value by a (configurable) amount. This expertise value could then be used for example to provide more explanatory information for novices

(on this particular topic), whereas experts could be served with more focussed compositions (i.e. containing less additional material on that particular topic).

Throughout the various interactions with such a system, users should be in full control of the fine-grained expertise level values in order to enable modifications of potential mismatches. Figure 6-20 presents an implementation of such a system, which uses the described interaction data in order to calculate and display fine-grained user expertise levels that can be modified manually through a slider interface.

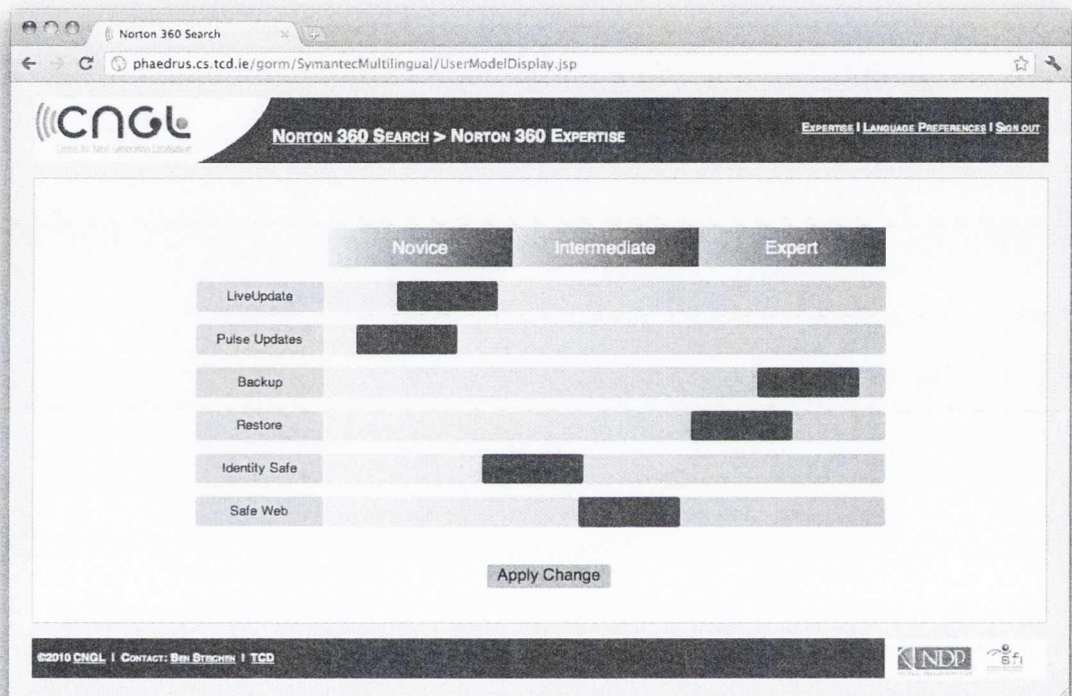


Figure 6-20. User expertise modelling

6.6.3. Multimedia support

Throughout this thesis, information compositions have mainly focussed on adaptively retrieving and composing textual-only information. However, as described in sections 4.3.1 and 4.3.2, the architecture design and implementation are also capable of adaptively retrieving and composing additional media types such as images or videos. Figure 6-21 shows an implementation of such a multimedia composition, where “Videos from the Web” are displayed alongside “Instructions”.

In particular, such multimedia capabilities could be adaptively used for specific tasks, contexts and preferences. For example, instructional videos could be retrieved and

composed specifically for queries requesting “how-to” information. Similarly, users who prefer highly visual material to fully understand certain concepts might be best served with a composition that contains textual results, as well as images and/or videos.

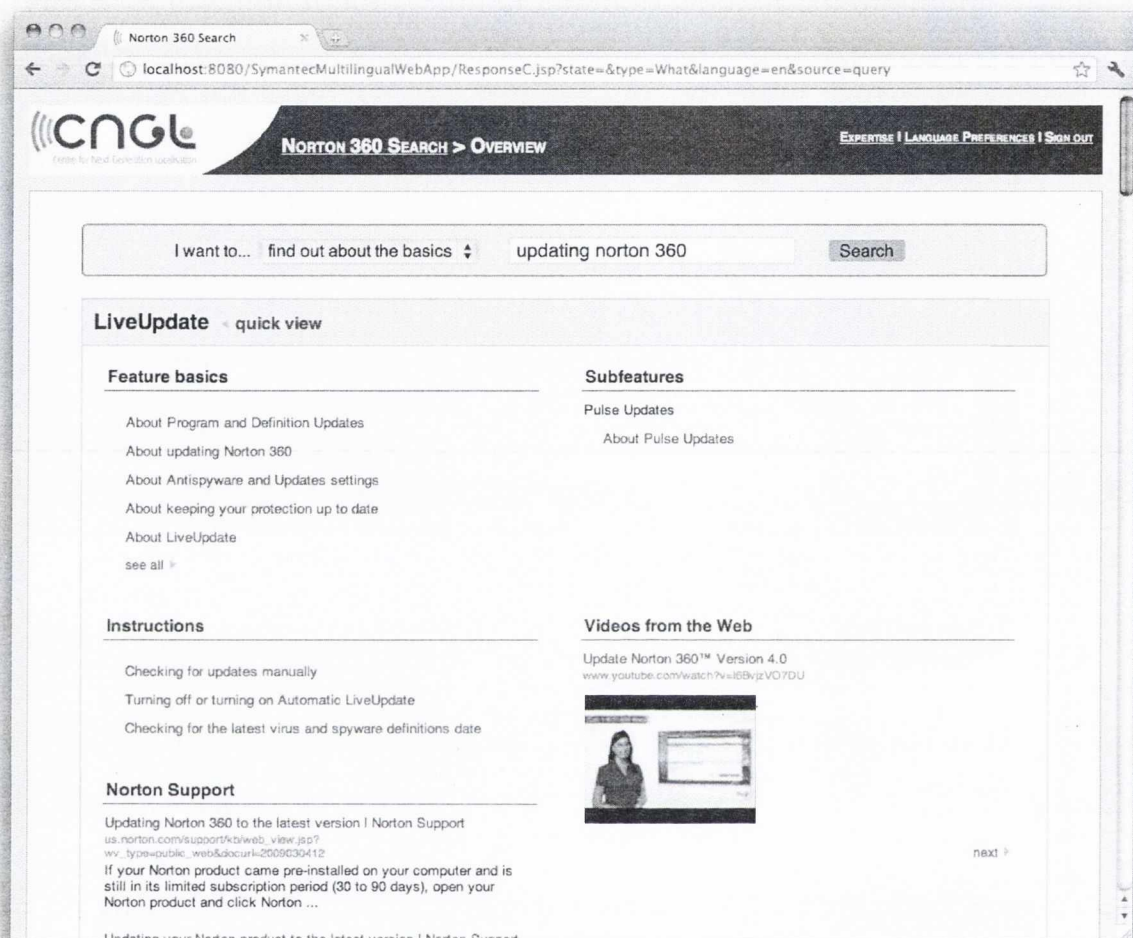


Figure 6-21. Information composition including video results

6.6.4. Alternative customer care application

Although the implementation prototypes presented above have each focussed on the Symantec product Norton 360, their underlying techniques and architectures are designed in an application-independent manner. In order to demonstrate this genericity, the techniques and architecture were also applied in an alternative application.

The objective of the alternative implementation was to evaluate if the approach could produce relevant information compositions for a different application/content area. More specifically, the alternative application aimed at testing if the infrastructure could cope with a completely different set of models and content. To this end, a prototype was developed for an alternative customer care application, which provided the same

functionalities as presented above, but focussing on Microsoft Office products such as Word and Excel and using different content bases (e.g. Microsoft documentation, forums) (see Figure 6-22).

In order to implement this prototype, the only major implementation changes consisted of replacing the underlying domain and content models, as well as the specification of targeted open-corpus support sites (e.g. Microsoft forums). The underlying architectures remained the same as with previous prototypes, proving that the approach could be applied successfully for alternative application scenarios without requiring conceptual changes.

This implementation was demonstrated during a series of CNGL exhibitions, including a showcase at Microsoft where visitors and employees informally evaluated the prototype. The informal feedback confirmed that the information compositions were well-integrated and relevant, hence providing further evidence for the genericity of the approach.

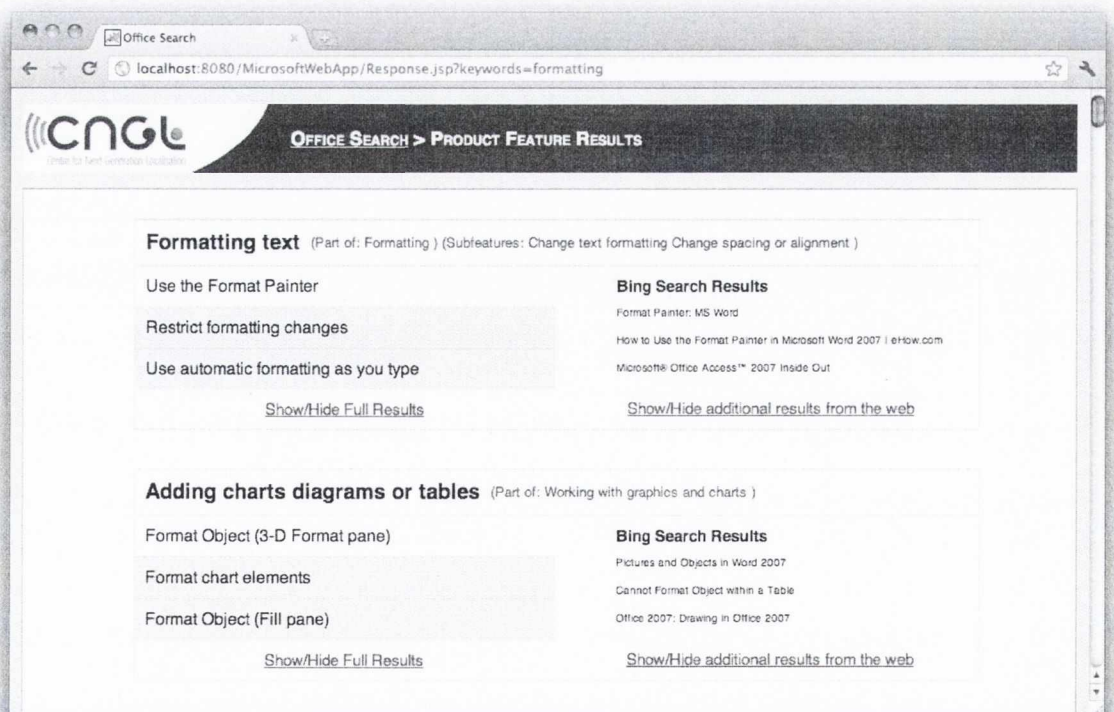


Figure 6-22. Microsoft Office prototype

6.7. Conclusions

Following the successful application and evaluation of the compositional approach presented in chapters 4 and 5, this chapter has described a range of additional adaptation capabilities that can be achieved using the presented approach, techniques and technologies.

First of all, the multilingual capabilities of ARCHING have been shown to successfully provide a novel approach for supporting users across a range of multilingual information sources. In particular, the evaluation results have revealed that bilingual users generally appreciate the novel approach of composing and presenting multilingual information in a fully integrated interface. It has been shown that the current paradigm of strict language separation can have negative effects on the overall user efficiency, effectiveness and satisfaction. Results have also shown that users are comfortable with more integrated compositions and that they do not feel overwhelmed by the amount and diversity of information. As with the monolingual evaluation, users generally found the most integrated and structured interface to provide increased task support and overall guidance towards personally relevant information.

Secondly, a number of additional supported adaptation dimensions have been presented, including mobile device adaptation, fine-grained user expertise modelling, as well as multimedia support. This additional support has confirmed the multidimensional capabilities of the approach and its transferability to diverse usage and content bases without the need for changes to the underlying architectures.

In conclusion, this chapter has confirmed that the adaptive retrieval, composition and presentation of closed-corpus and open-corpus information can be applied successfully in order to support authentic user information needs. Moreover, it has been shown that the techniques can be applied in a number of distinct implementations in order to support a multitude of user information needs, preferences, capabilities and contexts.

7 Conclusions

7.1. Introduction

This chapter concludes the thesis through a discussion of the overall achievements and contributions, as well as potential future research directions. In particular, section 7.2 reiterates the research question of the thesis and analyses how well the respective objectives have been achieved. Section 7.3 discusses the overall contributions of this thesis and presents the research publications that have resulted from this work. Finally, section 7.4 outlines a number of future directions for adaptive information composition research.

7.2. Research question and objectives

As stated in chapter 1 (section 1.2), *this thesis is researching the techniques and technologies required to generate adaptive information compositions that satisfy informal queries according to multiple dimensions of adaptation across closed-corpus and open-corpus information sources. More specifically, it asks the question, “what adaptive techniques and technologies are needed to provide such multidimensional information compositions across closed and open corpora in order to enhance a user’s effectiveness, efficiency and satisfaction.”*

In order to address this question, the thesis has proposed the notion of a compositional approach to information retrieval and delivery, which encompasses adaptive retrieval, composition and presentation in an integrated adaptation process. This approach has been achieved through a novel combination of techniques and technologies from the areas of Adaptive Hypermedia (AH) and Personalised Information Retrieval (PIR). Moreover, these techniques have been integrated in an architecture called ARCHING

(Adaptive Retrieval and Composition from Heterogeneous Information for personalised hypertext Generation), which is shown to be capable of generating adaptive information compositions that (i) satisfy informal queries (through the integration of closed- and open-corpus PIR capabilities) (ii) according to multiple dimensions of adaptation (through the integration of multi-model AH capabilities) and (iii) across closed-corpus and open-corpus information sources (through the combination of closed- and open-corpus manipulation capabilities in an integrated architecture).

A series of ARCHING-based prototypes have been developed and evaluated, which have shown that the compositional approach can significantly increase a user's effectiveness, efficiency and satisfaction in authentic information seeking tasks. Moreover it has been shown with evidence that the approach can adapt to a multitude of user dimensions (e.g. prior knowledge, query intent, user languages, device capabilities) and that different composition are particularly suited to individual user characteristics.

In addition to addressing the overall research question, this thesis has also met the individual research objectives that were identified in chapter 1 (section 1.3). Each of these objectives is discussed individually below.

Research objective (1): *Identify key affordances, techniques and impacts of current adaptive information access systems, particularly in the areas of Adaptive Hypermedia (AH) and Personalised Information Retrieval (PIR).*

This objective has been achieved through a comparative survey of AH and PIR techniques across three search process adaptation stages, namely *query adaptation*, *retrieval adaptation* and *adaptive composition and presentation* (chapter 2). For each of these stages, the survey has identified respective affordances and impacts through a structured analysis across a number of adaptation characteristics, including adaptation algorithms, model/metadata usage, scalability and adaptive behaviour.

In this analysis, the statistical, keyword-based algorithms of PIR (mainly to perform query expansion and retrieval adaptation) have been found to be highly scalable due to their low metadata requirements. However, it has also been revealed that such techniques generally only focus on user interests inferred from prior interactions (e.g. queries, clicks), thereby not addressing the array of additional user dimensions. Moreover, these techniques are typically confined to the generation of improved ranked

lists, thereby not providing any additional user guidance through adaptive composition, navigation or presentation support.

By contrast, AH techniques have been found to be inherently focussed on providing such user guidance through multiple models that enable particular information seeking strategies. Moreover, through the use of such multi-model adaptation techniques, AH systems also address a number of additional user dimensions (e.g. prior knowledge) through a multitude of adaptation techniques (e.g. adaptive composition, content annotation). However, the survey has also revealed that due the inherent reliance on refined concept and metadata indexing, most AH techniques have still been confined to very narrow domains, such as educational systems or cultural heritage libraries.

The conclusions of this comparative analysis of PIR and AH techniques have shown that the development of “hybridised” systems could potentially combine the open-corpus adaptation capabilities of PIR with the multidimensional adaptation capabilities of AH systems. Finally, it is argued that such a solution can address the research question by generatating multidimensional information compositions across closed-corpus and open-corpus content.

The achievement of research objective 1 is also presented in the following publication.

Steichen, B., Ashman, H., and Wade, V. (2012). A comparative survey of Personalised Information Retrieval and Adaptive Hypermedia techniques. *Information Processing and Management*, Volume 48, Issue 4, pp. 698–724.

Research objective (2): *Design and develop system architectures and adaptation processes that enable the generation of adaptive information compositions according to multiple levels of adaptation across closed- and open-corpus information sources.*

In order to address this objective, an initial adaptive open-corpus composition prototype was developed using a multi-model metadata-driven architecture (see chapter 3). This (educational) prototype was evaluated in terms of its ability to include open-corpus information, as well as its potential usability benefits and drawbacks in an authentic e-learning scenario. Evaluation findings showed that the architecture could successfully deliver information presentations that were entireley composed from open-corpus content. Moreover, it was shown that users recognised and valued the adaptive features and that they were motivated to browse and view more information compared to

conventional search systems. However, the presented (crowd-sourced) metadata generation process confirmed the difficulties and costs of integrating open-corpus information into such a system. Moreover, users were restricted during the query elicitation stage, as only concepts known to the system could be used as query inputs.

Motivated by the analysis of techniques and technologies of PIR and AH, as well as the findings of this first experiment, a more advanced approach was proposed, which combined closed-corpus and open-corpus adaptation capabilities in an integrated architecture called ARCHING (Adaptive Retrieval and Composition from Heterogeneous Information for personalised hypertext Generation) (see chapter 4). This architecture addresses all aspects of the research objective, as it enables the generation of adaptive information compositions according to multiple levels of adaptation (using the multi-model adaptation capabilities) across closed-corpus and open-corpus information sources (by combining closed- and open-corpus capabilities in an integrated architecture). Moreover, this architecture supports informal keyword queries, while still generating an adaptive information composition as a response. The architecture thereby fully retains the Adaptive Hypermedia capabilities of the prototype used in the first experiment, while integrating adaptive open-corpus modules that loosen up the strong reliance on metadata annotations.

In order to evaluate this approach and architecture, a prototype implementation was developed for a personalised customer care scenario. This case study provided real-life information needs, closed- and open-corpus content, as well as authentic evaluation possibilities. The evaluation of the ARCHING-based prototype has shown that the research objective has been met successfully, as users were able to solve authentic customer care tasks using information compositions that were adaptively composed from closed-corpus and open-corpus information sources. Moreover, through the development of a series of additional prototypes, it has been shown that such compositions can be adapted and applied to multiple user dimensions (see chapters 5 and 6).

The achievement of research objective 2 is also presented in the following publications.

Steichen, B., Lawless, S., O'Connor, A., and Wade, V. (2009). Dynamic hypertext generation for reusing open corpus content. *In Proceedings of the 20th ACM conference on Hypertext and hypermedia (HT '09)*, pp. 119-128.

Steichen, B., and Wade, V. Adaptive Retrieval and Composition of Socio-Semantic Content for Personalised Customer Care. (2010). *In Proceedings of the International Workshop on Adaptation in Social and Semantic Web (SAS-WEB 2010), in conjunction with UMAP 2010*, CEUR Workshop Proceedings, ISSN 1613-0073, pp. 1-10.

Steichen, B., O'Connor, A., and Wade, V. (2011). Personalisation in the Wild - Providing Personalisation across Semantic, Social and Open-Web Resources. *In Proceedings of the 22nd ACM conference on Hypertext and hypermedia (HT '11)*, pp. 73-82.

Research objective (3): *Evaluate the architectures through a series of case-study implementations using metrics related to user efficiency, effectiveness and satisfaction.*

This objective has been achieved through a series of prototype implementations that have each been evaluated in terms of their respective usability. More specifically, in order to measure the user efficiency, effectiveness and satisfaction, each prototype implementation has been evaluated using authentic task-based user studies. A multitude of metrics have been used throughout these evaluations, including task times, number of queries, page views and task success (to measure user efficiency and effectiveness), as well as perceived usability scores, task confidence and perceived assistance and guidance (to measure user satisfaction).

The initial (educational) prototype has revealed that information compositions can successfully support users in an e-learning environment (i.e. help students achieve a significant knowledge gain), and that users are motivated to browse and view more information (see section 3.5). Moreover, the evaluation has shown that users appreciate the result relevance, presentation and composition of the adaptive prototype.

In contrast to this educational evaluation, the second evaluation has focussed on measuring the user efficiency, effectiveness and satisfaction in a customer support scenario, where it is crucial to allow users to find relevant information as quickly as possible (while still maintaining the feeling of personalised guidance and task assistance) (see section 4.5). The results of this evaluation have shown that users issue significantly fewer queries and complete their tasks significantly faster with an adaptive composition system compared to a non-adaptive search system. Moreover, users are still able to view more information and generally feel that the adaptive system provides

more personally relevant information and task assistance. In addition, a range of user satisfaction questionnaires have revealed that users significantly prefer the concept of adaptive information compositions compared to conventional search systems. A caveat has to be noted that this evaluation consisted of only 36 participants. However, the results and opinions were further confirmed in subsequent evaluations with over 120 users.

In a third round of evaluations, a range of distinct (monolingual and multilingual) prototypes have been evaluated in terms of their comparative user efficiency, effectiveness and satisfaction (see sections 5.5 and 6.5). Results have confirmed that the concept of adaptive information compositions can successfully support users in open-web information seeking tasks. Moreover, this evaluation has revealed that users' efficiency and effectiveness vary with different composition systems and that the respective user satisfaction depends on individual user preferences and characteristics.

In total, the various prototype implementations have been evaluated by almost 200 participants in (anonymous) online-based user studies. These experiments have each confirmed that the compositional approach to information retrieval and delivery can successfully provide user efficiency, effectiveness and satisfaction in authentic real-life application scenarios. Moreover, it has been shown that a multitude of user preferences, characteristics and contexts can be supported in order to provide the most personally suitable information at the right time and in the right form.

The achievement of research objective 3 is also presented in the following publications.

Steichen, B., Lawless, S., O'Connor, A., and Wade, V. (2009). Dynamic hypertext generation for reusing open corpus content. *In Proceedings of the 20th ACM conference on Hypertext and hypermedia (HT '09)*, pp. 119-128.

Steichen, B., and Wade, V. Adaptive Retrieval and Composition of Socio-Semantic Content for Personalised Customer Care. (2010). *In Proceedings of the International Workshop on Adaptation in Social and Semantic Web (SAS-WEB 2010), in conjunction with UMAP 2010*, CEUR Workshop Proceedings, ISSN 1613-0073, pp. 1-10.

Steichen, B., O'Connor, A., and Wade, V. (2011). Personalisation in the Wild - Providing Personalisation across Semantic, Social and Open-Web

Resources. In *Proceedings of the 22nd ACM conference on Hypertext and hypermedia (HT '11)*, pp. 73-82.

7.3. Contributions

The major contribution of this thesis is a novel *compositional approach to open- and closed-corpus mono- and multilingual* information retrieval and delivery, which combines Adaptive Hypermedia and Personalised Information Retrieval techniques. In particular, this innovative approach is capable of retrieving and composing closed-corpus and open-corpus content in order to generate adaptive information presentations according to multiple dimensions of adaptation. Moreover, this approach enables the integration of adaptive navigation and presentation techniques across mono- and multilingual closed- and open-corpus information. Through a series of case-study implementations, this approach has been shown to provide a range of user benefits in terms of user efficiency, effectiveness and satisfaction. This approach thereby advances the field of Adaptive Hypermedia by providing a novel, metadata-sparse approach to include open-corpus resources. Moreover, the approach advances the field of Personalised Information Retrieval by providing a range of novel presentation and adaptation techniques in order to support multiple user dimensions.

In addition, the thesis has presented an architecture called ARCHING (Adaptive Retrieval and Composition from Heterogeneous Information for personalised hypertext Generation), which implements this approach by combining multi-model Adaptive Hypermedia capabilities with adaptive open-corpus manipulation techniques. This architecture has been used as the basis for a series of prototypes, which have demonstrated the successful application of the compositional approach for authentic information seeking tasks. The adaptive prototypes have been shown to successfully retrieve and compose closed-corpus and open-corpus information sources in order to provide multidimensional user assistance and guidance. Through this demonstration of benefits for using a compositional approach, the thesis has presented new directions for the development of novel adaptive information access systems. In particular, it has been shown that prototypes building on this multidimensional approach can successfully support users through information composition presentations that go beyond the current one-dimensional, ranked-list based information retrieval paradigm.

The contributions of this research have also resulted in a number of high-quality scientific publications, which will be discussed next.

The analysis of Personalised Information Retrieval and Adaptive Hypermedia techniques (presented in chapter 2) has been published in the journal of *Information Processing and Management*. In particular, this paper presents a review of current techniques and technologies in both fields and argues for the development of hybrid information access systems in order to leverage their complementary affordances.

Steichen, B., Ashman, H., and Wade, V. (2012). A comparative survey of Personalised Information Retrieval and Adaptive Hypermedia techniques. *Information Processing and Management*. Volume 48, Issue 4, pp. 698–724.

The design, development and evaluation of the initial information composition prototype (presented in chapter 3) have been published in a full conference paper at the *20th ACM conference on Hypertext and hypermedia (HT '09)*. This paper demonstrates the successful application of the compositional approach across externally-sourced open-corpus information and presents the metadata requirements for this initial prototype.

Steichen, B., Lawless, S., O'Connor, A., and Wade, V. (2009). Dynamic hypertext generation for reusing open corpus content. *In Proceedings of the 20th ACM conference on Hypertext and hypermedia (HT '09)*, pp. 119-128.

The ARCHING architecture (presented in chapter 4) and its application potential in a personalised customer care scenario have been published at the *International Workshop on Adaptation in Social and Semantic Web (SAS-WEB 2010)*, which was held in conjunction with the *18th conference on User Modeling, Adaptation, and Personalization (UMAP 2010)*. In particular, this paper presents the ability of the architecture to retrieve, compose and present information from semantic (closed-corpus), as well as social (open-corpus) information sources.

Steichen, B., and Wade, V. Adaptive Retrieval and Composition of Socio-Semantic Content for Personalised Customer Care. (2010). *In Proceedings of the International Workshop on Adaptation in Social and Semantic Web (SAS-WEB 2010), in conjunction with UMAP 2010*, CEUR Workshop Proceedings, ISSN 1613-0073, pp. 1-10.

The complete design, development and evaluation of the personalised customer care prototype (presented in chapter 4) have been published in a full conference paper at the *22nd ACM conference on Hypertext and hypermedia (HT '11)*. In particular, this paper describes the successful application of ARCHING to support authentic user needs and presents the resulting increase in user efficiency, effectiveness and satisfaction compared to a purpose-built baseline system.

Steichen, B., O'Connor, A., and Wade, V. (2011). Personalisation in the Wild - Providing Personalisation across Semantic, Social and Open-Web Resources. *In Proceedings of the 22nd ACM conference on Hypertext and hypermedia (HT '11)*, pp. 73-82.

In addition to the scientific contributions of this thesis, the presented research work has also attracted significant commercial interest. In particular, this work has been selected for the Enterprise Ireland Innovation Partnership program⁵⁷, which aims to commercialise the presented techniques and technologies in a spin-out company for personalised multilingual customer care⁵⁸. Moreover, this work has received commercialisation funding from the Science Foundation Ireland⁵⁹, which provides additional resources to the research and development of the presented techniques and technologies.

Technology and Innovation Development Award (TIDA) Feasibility 2011. *An integrated software suite to provide Next-generation Personalised Multilingual Customer Care*. (SFI 11/TIDA/I1985).

7.4. Future Work

The research described in this thesis has resulted in a novel compositional approach to information retrieval and delivery, which has been proven successful through several case study applications. Building upon these results, there are a number of research opportunities for future work.

⁵⁷ <http://www.enterprise-ireland.com/en/Funding-Supports/Researcher/Funding-to-Collaborate-with-Industry-in-Ireland/Innovation-Partnerships.shortcut.html>

⁵⁸ <http://emizar.com/>

⁵⁹ <http://www.sfi.ie/>

Additional applications domains

The techniques and technologies presented in this thesis have been inherently conceived to be application- and domain-independent. In particular, the closed-corpus multi-model adaptation techniques, as well as the open-corpus manipulation techniques can be reused flexibly without changing the underlying architecture. More specifically, since the various models are separated from the actual adaptation, retrieval and presentation capabilities, it is possible to apply the presented techniques to a multitude of alternative application domains.

For example, the latest version of the ARCHING architecture could be used for an enhanced, more open educational support system. Such a system could include closed-corpus educational resources and teaching strategies from an educator, as well as open-web resources to increase the choice and diversity of information presented to the student.

In the initial SQL scenario (presented in chapter 3), such an open-web system could for example adaptively integrate additional open-web educational resources (e.g. w3schools⁶⁰), as well as open-corpus resources from a number of different SQL database vendors (e.g. MySQL⁶¹ or Oracle⁶²). By maintaining the adaptive composition and guidance across these various information sources, it would therefore be possible to ensure the educational coherence of the presented information without restricting the overall scope of the material.

Further example applications for the compositional approach could be in commercial domains such as e-tourism or e-commerce. As with the presented customer care application domain, users could be guided across a range of close-corpus and open-corpus information sources in order to receive a more informative and coherent search experience. In particular, due to the high page view count that was measured throughout the information composition evaluations, information providers may achieve a higher *website stickiness*⁶³, which could ultimately lead to increased sales revenues.

⁶⁰ <http://www.w3schools.com>

⁶¹ <http://www.mysql.com>

⁶² <http://www.oracle.com>

⁶³ http://en.wikipedia.org/wiki/Sticky_content

Multilingual Information Compositions

As part of the research presented in this thesis, a novel approach to multilingual information access has been proposed, which is tailored towards a user's individual language preferences and proficiencies. An initial evaluation has presented clear benefits of such an approach for bilingual users, with participants significantly preferring multilingual result presentations to the more conventional language separation.

These initial results open up a number of significant research opportunities in the development of novel multilingual information access systems, especially considering the different degrees of language proficiencies of web users. While much emphasis in current cross-lingual information retrieval research has been placed on improving translation and retrieval effectiveness, there has been a clear research gap in the area of multilingual interfaces. The presented adaptive compositional approach allows a multitude of novel composition and presentation possibilities, which can be used to research and address this pressing issue of adaptive, multilingual search interaction.

Task-based Evaluation Framework

The evaluation methodology used in this thesis has consisted of task-based experiments, whereby research prototypes have been assessed through comparative user studies. By placing the user at the centre of the evaluation, this methodology has enabled the evaluation of the real-world applicability of the proposed approaches and techniques. However, this type of evaluation has been found to constitute a very expensive process in terms of development time and recruitment effort.

To this end, the experiments presented in this thesis have been carried out using an evaluation framework that allowed users to participate online, using their own equipment and in their own time. This enabled the recruitment of large numbers of participants, as well as the provision of the most realistic evaluation scenario (since users did not experience an unknown environment).

A significant contribution to the research community would therefore lie in the development of a generic evaluation framework, which could be offered to information access researchers as a configurable, task-based experiment platform. This may also be

coupled with an online portal where common evaluation methods and metrics could be shared across researchers. Moreover, such a portal could include “call for evaluation participation” noticeboards in order to provide a common platform for the recruitment of participants.

Bibliography

- Abel, F., Gao, Q., Houben, G.J., and Tao, K. (2011). Semantic Enrichment of Twitter Posts for User Profile Construction. *Proceedings of 8th Extended Semantic Web Conference (ESWC)*, pp. 375-389.
- Ahn, J.-W., Brusilovsky, B., Grady, J., He, D. and Florian, R. (2010). Semantic annotation based exploratory search for information analysts. *Information Processing and Management*, 46, pp. 383-402.
- Aroyo, L., De Bra, P., and Houben, G.J. (2003). Embedding Information Retrieval in Adaptive Hypermedia: IR meets AHA!. *Proceedings of the Workshop on Adaptive Hypermedia and Adaptive Web-Based Systems, at the 12th World Wide Web Conference, WWW2003*, pp. 63-76.
- Ashman, H., Brailsford, T. and Brusilovsky, P. (2009). Personal Services: Debating the Wisdom of Personalisation. *In Proceedings of the 8th International Conference on Advances in Web Based Learning (ICWL '09)*, pp. 1-11.
- Bakalov, F., König-Ries, B., Nauerz, A., and Welsch, M. (2010) Introspective views: An interface for scrutinizing semantic user models. *Proceeding of the 18th International Conference on User Modeling, Adaptation, and Personalization, UMAP 2010*, pp. 219-230.
- Benedek, J., and Miner, T. (2002). Product Reaction Cards. <http://www.microsoft.com/usability/UEPostings/ProductReactionCards.doc>. Developed by and © 2002 Microsoft Corporation. All rights reserved.
- Berners-Lee, T., Hendler, J., and Lassila O. (2001). The Semantic Web. *Scientific American*, 284, 5, pp.35-43.
- Bhavnani, S., Bichakjian, C., Johnson, T., Little, R., Peck, F., Schwartz, J., and Strecher, V. (2003) Strategy hubs: next-generation domain portals with search procedures. *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 393-400.
- Bhagal, J., Macfarlane, A., and Smith, P. (2007). A review of ontology based query expansion. *Information Processing and Management*. 43, 4, pp. 866-886.

- Billsus, D., and Pazzani, M. J. (2007). Adaptive News Access. *The Adaptive Web*, LNCS, vol. 4321, pp. 550-570.
- Brin, S., and Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30, 1-7, pp. 107-117.
- Brooke, J. (1996). SUS: a "quick and dirty" usability scale. Usability Evaluation in Industry. *London: Taylor and Francis*.
- Brown, P. J., and Jones, G. J. F. (2001). Context-aware Retrieval: Exploring a New Environment for Information Retrieval and Information Filtering. *Personal Ubiquitous Computing*, 5, 4, pp. 253-263.
- Brusilovsky, P., and Henze, N. (2007). Open Corpus Adaptive Educational Hypermedia. *The Adaptive Web*, LNCS, vol. 4321, pp. 671-696.
- Brusilovsky, P., Chavan, G., and Farzan, R. (2004). Social adaptive navigation support for open corpus electronic textbooks. *Adaptive Hypermedia and Adaptive Web-Based Systems, AH 2004*, LNCS, vol. 3137, pp. 24-33.
- Brusilovsky, P., Kobsa, A., and Nejdl, W. (2007). *The Adaptive Web: Methods and Strategies of Web Personalization*, LNCS, Vol. 4321, Preface.
- Burke, R. (2007). Hybrid web recommender systems. In *The Adaptive Web: Methods and Strategies of Web Personalization*, LNCS, vol. 4321, pp. 377-408.
- Burton-Jones, A., Storey, V., Sugumaran, V., and Sandeep, P. (2003). A heuristic-based methodology for semantic augmentation of user queries on the Web. *Conceptual Modeling - ER 2003*, LNCS, vol. 2813, pp. 476-489.
- Calegari, S., and Pasi, G. Personalized Ontology-Based Query Expansion. (2008). *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, vol. 3, pp. 256-259.
- Cantador, I., Fernández, M., Vallet, D., Castells, P., Picault, J., and Ribière, M. (2008). A Multi-Purpose Ontology-Based Approach for Personalised Content Filtering and Retrieval. *Advances in Semantic Media Adaptation and Personalization. Springer-Verlag, Studies in Computational Intelligence*, vol. 93, pp. 25-51.

- Carpineto, C., Mori, R.D., Romano, G., and Bigi, B. (2001). An information-theoretic approach to automatic query expansion. *ACM Transactions on Information Systems*, 19, 1, pp. 1-27.
- Chirita, P., Firan, C. S., and Nejdl, W. Personalized query expansion for the web. (2007). *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '07*, pp. 7-14.
- Conlan, O., Wade, V., Bruen, C. and Gargan, M. (2002). Multi-model, Metadata Driven Approach to Adaptive Hypermedia Services for Personalized eLearning. *Adaptive Hypermedia and Adaptive Web-Based Systems, AH2002*, LNCS, vol. 2347, pp. 100-111.
- Conlan, O., and Wade, V. (2004). Evaluation of APeLS – An Adaptive eLearning Service Based on the Multi-model, Metadata-Driven Approach. *Adaptive Hypermedia and Adaptive Web-Based Systems, AH2004*, LNCS, vol. 3137, pp. 291-295.
- Croft, B., Metzler, D. and Strohman, T. (2009). Search Engines: Information Retrieval in Practice (1st ed.). *Addison-Wesley Publishing Company*, USA.
- Daoud, M., Tamine, L., and Boughanem, M. (2010). A personalized graph-based document ranking model using a semantic user profile. *Proceedings of the 18th international conference on User modeling, adaption, and personalization, UMAP 2010*, pp. 171–182.
- De Bra, P., Aerts, A., Berden, B., de Lange, B., Rousseau, B., Santic, T., Smits, D., and Stash, N. (2003). AHA! The adaptive hypermedia architecture. *Proceedings of the Fourteenth ACM Conference on Hypertext and Hypermedia, HYPERTEXT '03*, pp. 81-84.
- Dey, A. (2001). Understanding and Using Context. *Personal and Ubiquitous Computing*, 5, 1, pp. 4–7.
- Dolog, P., Henze, N., Nejdl, W. and Sintek, M. (2003). Towards the adaptive Semantic Web. *Principles and Practice of Semantic Web Reasoning*, LNCS, vol. 2901, pp. 51-68.

- Farrell, R. G., Liburd, S. D., and Thomas, J. C. (2004). Dynamic assembly of learning objects. *Proceedings of the 13th international World Wide Web Conference on Alternate Track Papers & Posters, WWW Alt. '04*, pp. 162-169.
- Felder, R.M., and Silverman, L.K. (1988). Learning and Teaching Styles in Engineering Education, *Engineering Education*, 78(7), pp. 674-681.
- Fernandez, M., Lopez, V., Sabou, M., Uren, V., Vallet, D., Motta, E., and Castells, P. (2008). Semantic Search Meets the Web. *Proceedings of the 2008 IEEE international Conference on Semantic Computing*, pp. 253-260.
- Fu B., Brennan R., and O'Sullivan D. (2010) Cross-Lingual Ontology Mapping and Its Use on the Multilingual Semantic Web. *In Proceedings of the 1st Workshop on the Multilingual Semantic Web, at the 19th International World Wide Web Conference (WWW 2010)*, CEUR Vol.571, pp.13-20.
- Geurts, J., Bocconi, S., and Van Ossenbruggen, J. (2003). Towards ontology-driven discourse: From semantic graphs to multimedia presentations. *The Semantic Web - ISWC 2003*, LNCS, vol. 2870, pp. 597-612.
- Glover, E. J., Lawrence, S., Gordon, M. D., Birmingham, W. P., and Giles, C. L. (2001). Web Search-Your Way. *Communications of the ACM*, 44, 12, pp. 97-102.
- Haveliwala, T. H. (2003). Topic-Sensitive PageRank: A Context-Sensitive Ranking Algorithm for Web Search. *IEEE Transactions on Knowledge and Data Engineering*, 15, 4, pp. 784-796.
- Henze, N. (2005). Personal Readers: Personalized Learning Object Readers for the Semantic Web. *Proceedings of 12th International Conference on Artificial Intelligence in Education, AIED '2005*, pp. 274-281.
- Henze, N., and Nejd, W. (2001). Adaptation in open corpus hypermedia. *International Journal of Artificial Intelligence in Education*, 12, 4, pp. 325-350.
- Hirschman, L., and Gaizauskas, R. (2001). Natural language question answering: the view from here. *Natural Language Engineering*, 7, 4, pp. 275-300.

- Hsiao, I.-H., Brusilovsky, P., Yudelson, M., and Ortigosa, A. (2010). The value of adaptive link annotation in e-learning: a study of a portal-based approach. *Proceedings of the 21st ACM conference on Hypertext and hypermedia, HT '10*, pp. 223-228.
- Hsiao, I.-H., Sosnovsky, S., and Brusilovsky, P. (2009). Adaptive Navigation Support for Parameterized Questions in Object-Oriented Programming. *4th European Conference on Technology Enhanced Learning, EC-TEL 2009*, pp. 88-98.
- Jansen, B. J., Spink, A., and Saracevic, T. (2000). Real life, real users, and real needs: a study and analysis of user queries on the web. *Information Processing and Management*. 36, 2, pp. 207-227.
- Jansen, B. J., Booth, D. L., and Spink, A. (2008). Determining the informational, navigational, and transactional intent of Web queries. *Information Processing and Management*. 44, 3, pp. 1251-1266.
- Joho, H., and Jose, J. M. (2008). Effectiveness of additional representations for the search result presentation on the web. *Information Processing and Management*, 44, 1, pp. 226-241.
- Jovanović, J., Gašević, D., and Devedžić, V. (2006). Dynamic Assembly of Personalized Learning Content on the Semantic Web. *The Semantic Web: Research and Applications*, LNCS, vol. 4011, pp. 545-559.
- Karam, N., Linckels, S. and Meinel, C. (2007). Semantic composition of lecture subparts for a personalized e-learning. *The Semantic Web: Research and Applications*, LNCS, vol. 4519, pp. 716-728.
- Kim, M., and Choi, K. (1999). A comparison of collocation-based similarity measures in query expansion. *Information Processing and Management*, 35, 1, pp. 19-30.
- Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46, 5, pp. 604-632.
- Koutrika, G., Ioannidis, Y. (2005). A Unified User Profile Framework for Query Disambiguation and Personalization. *Proceedings of the Workshop on New Technologies for Personalized Information Access, PIA2005*, pp. 44-53.

- Krovetz, R., (1997). Homonymy and Polysemy in Information Retrieval. *Proceedings of the Eighth Conference on European Chapter of the Association for Computational Linguistics*, pp. 72–79.
- Kumaran, G. and Allan, J. (2008). Effective and efficient user interaction for long queries. *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '08*, pp. 11-18.
- Lawless, S., Hederman, L., and Wade, V. Enhancing Access to Open Corpus Educational Content: Learning in the Wild. *Proceedings of the 19th ACM Conference on Hypertext and Hypermedia, Hypertext 2008*, pp. 167-174.
- Lin, Y.-l., and Brusilovsky, P. (2011). Towards open corpus adaptive hypermedia: a study of novelty detection approaches. *Proceedings of the 19th international conference on User modeling, adaption, and personalization, UMAP'11*, pp. 353-358.
- Linckels, S., Sack, H., and Meinel, C. (2007). Optimizing the retrieval of pertinent answers for nl questions with the e-librarian service. *Service Matchmaking and Resource Retrieval in the Semantic Web, SMR2*, pp. 139-152.
- Liu, F., Yu, C., and Meng, W. (2004). Personalized Web Search For Improving Retrieval Effectiveness. *IEEE Transactions on Knowledge and Data Engineering*, 16, 1, pp. 28–40.
- Mangold, C. (2007). A survey and classification of semantic search approaches. *International Journal of Metadata, Semantics and Ontologies*, 2, 1, pp. 23-34.
- Manning, C.D., Raghavan, P., and Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.
- Marchionini, G. (2006). Exploratory search: from finding to understanding. *Communications of the ACM*, 49, 4, pp. 41-46.
- Maekawa, T., Yanagisawa, Y., Sakurai, Y., Kishino, Y., Kamei, K. and Okadome, T. (2009) Web searching for daily living. In *Proceedings of the 32nd*

international ACM SIGIR conference on Research and development in information retrieval (SIGIR '09), pp. 27-34.

Micarelli, A., Gasparetti, F., and Biancalana, C. (2006). Intelligent search on the Internet. *Reasoning, Action and Interaction in AI Theories and Systems*, pp. 247–264.

Micarelli, A., and Sciarrone F. (2004). Anatomy and empirical evaluation of an adaptive web-based information filtering system. *User Modeling and User-Adapted Interaction*, 14, 2-3, pp. 159-200.

Miller, G. (1995). WordNet: a lexical database for English. *Communications of the ACM*, 38, 11, pp. 39-41.

Navigli, R., and Velardi, P. (2003). An analysis of ontology-based query expansion strategies. *Workshop on adaptive text extraction and mining, ATEM 2003, at the 14th European conference on machine learning, ECML 2003*, pp. 42-49.

Noh, H.-Y., Lee, J.-H., Oh, S.-W., Hwang, K.-S., and Cho, S.-B. (2011). Exploiting indoor location and mobile information for context-awareness service. *Information Processing and Management*, 48, 1, pp. 1-12.

Oard, D.W., and Diekema, A.R. (1998). Cross-Language Information Retrieval. *Annual Review of Information Science (ARIST)*, pp. 223-256.

Oard, D.W. (2009). Multilingual Information Access. In *Encyclopedia of Library and Information Sciences*, 3rd Ed., Taylor & Francis.

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. New York, NY: Penguin.

Pitkow, J., Schütze, H., Cass, T., Cooley, R., Turnbull, D., Edmonds, A., Adar, E., and Breuel, T. (2002). Personalized search. *Communications of the ACM*, 45, 9, pp. 50-55.

Psarras, I., and Jose, J. (2006). A System for Adaptive Information Retrieval. *Adaptive Hypermedia and Adaptive Web-Based Systems, AH2006*, LNCS, vol. 4018, pp. 313-317.

- Radlinski, F. and Dumais, S. (2006). Improving personalized web search using result diversification. *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '06*, pp. 691-692.
- Reeve, L., and Han, H. (2005). Survey of semantic annotation platforms. *Proceedings of the 2005 ACM symposium on Applied computing (SAC '05)*, pp. 1634-1638.
- Reichheld, F. (2003). The One Number You Need. In: Harvard Business Review, December 2003.
- Rocchio, J. J. (1971). Relevance feedback in information retrieval. In G. Salton (Ed.), *The SMART retrieval system. Experiments in automatic document processing*, pp. 313-323. Englewood Cliffs, NJ: Prentice-Hall.
- Rocha, C., Schwabe, D., and Aragao, M. P. (2004). A hybrid approach for searching in the Semantic Web. *Proceedings of the 13th international Conference on World Wide Web, WWW '04*, pp. 374-383.
- Sah, M., and Wade, V. (2010). Automatic Metadata Extraction from Multilingual Enterprise Content. *Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM 2010)*, pp. 1665-1668.
- Sah, M., and Wade, V. (2011). Automatic mining of cognitive metadata using fuzzy inference. *Proceedings of the 22nd ACM conference on Hypertext and hypermedia, HT '11*, pp. 37-46.
- Schafer, J. B., Frankowski, D., Herlocker, J., and Sen, S. (2007). Collaborative filtering recommender systems. In *The Adaptive Web: Methods and Strategies of Web Personalization*, LNCS, vol. 4321, pp. 291-324.
- Sintek, M, and S Decker. (2002). TRIPLE-A Query, Inference, and Transformation Language for the Semantic Web. *The Semantic Web - ISWC 2002*, LNCS, vol. 2342, pp. 364-378.
- Smith, A.S.G., and Blandford, A. (2003). MLTutor: An application of machine learning algorithms for an adaptive Web-based information system. *International Journal of Artificial Intelligence in Education*, 13, 2-4, pp. 235-261.

- Smits, D., and De Bra, P. (2011). GALE: a highly extensible adaptive hypermedia engine. *Proceedings of the 22nd ACM conference on Hypertext and hypermedia, HT '11*, pp. 63-72.
- Speretta, M., and Gauch, S. (2005). Personalized Search Based on User Search Histories. *Web Intelligence, WI2005*, pp. 622–628.
- Spink, A., and Jansen, B.J. (2004). A study of web search trends. *Webology* 1, 2. <http://www.webology.ir/2004/v1n2/a4.html>.
- Stamou, S., and Ntoulas, A. (2009). Search personalization through query and page topical analysis. *User Modeling and User-Adapted Interaction*, 19, 1-2, pp. 5-33.
- Steichen, B., Lawless, S., O'Connor, A., and Wade, V. (2009). Dynamic hypertext generation for reusing open corpus content. In *Proceedings of the 20th ACM conference on Hypertext and hypermedia (HT '09)*, pp. 119-128.
- Steichen, B., and Wade, V. (2010). Adaptive Retrieval and Composition of Socio-Semantic Content for Personalised Customer Care. In *Proceedings of the International Workshop on Adaptation in Social and Semantic Web (SAS-WEB 2010)*, in conjunction with the 19th conference on User Modeling, Adaptation and Personalization (UMAP 2010), *CEUR Workshop Proceedings*, ISSN 1613-0073, pp. 1-10.
- Steichen, B., O'Connor, A., and Wade, V. (2011). Personalisation in the Wild - Providing Personalisation across Semantic, Social and Open-Web Resources. In *Proceedings of the 22nd ACM conference on Hypertext and hypermedia (HT '11)*, pp. 73-82.
- Steichen, B., Ashman, H., and Wade, V. (2012). A comparative survey of Personalised Information Retrieval and Adaptive Hypermedia techniques. *Information Processing & Management*, Volume 48, Issue 4, pp. 698–724.
- Sushmita, S., Joho, H., Lalmas, M. and Villa, R. (2010). Factors affecting click-through behavior in aggregated search interfaces. *Proceedings of the 19th ACM international conference on Information and knowledge management, CIKM '10*, pp. 519-528.

- Tanudjaja, F., and Mui, L. (2002). Persona: A Contextualized and Personalized Web Search. *Proceedings of the 35th Annual Hawaii international Conference on System Sciences, HICSS '02*, 3, 3, pp. 67.
- Teevan, J., Dumais, S. T., and Horvitz, E. (2005). Personalizing search via automated analysis of interests and activities. *Proceedings of the 28th international ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '05*, pp. 449-456.
- Teevan, J., Dumais, S., and Gutt, Z. (2008). Challenges for supporting faceted search in large, heterogeneous corpora like the Web. *Proceedings of HCIR 2008*, pp. 6-8.
- Thomas, P., Noack, K., and Paris, C. (2010). Evaluating interfaces for government metasearch. In *Proceeding of the third symposium on Information interaction in context (IiX '10)*, pp. 65-74.
- Tran, T., Wang, H., Lamparter, S., and Cimiano, P. (2008). Personalization Using Ontologies and Rules. *Adaptive Hypermedia and Adaptive Web-Based Systems, AH2008, LNCS*, vol. 5149, pp. 349-352.
- Truran, M., Goulding, J., and Ashman, H. (2005). Co-active intelligence for image retrieval. *Proceedings of the 13th Annual ACM international Conference on Multimedia*, pp. 547-550.
- Tsandilas, T. and schraefel, m.c. (2004). Usable adaptive hypermedia systems. *New Review of Hypermedia and Multimedia*, 10, 1, pp. 5-29.
- Tvarozek, M. and Bielikova, M. (2007). Adaptive faceted browser for navigation in open information spaces. *Proceedings of the 16th international Conference on World Wide Web, WWW '07*, pp. 1311-1312.
- Villa, R., Cantador, I., Joho, H., and Jose, J. M. (2009). An aspectual interface for supporting complex search tasks. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval (SIGIR '09)*, pp. 379-386

- Wade, V. (2009). Challenges for the Multi-Dimensional Personalised Web. *Keynote at 17th International Conference on User Modeling, Adaptation, and Personalization, UMAP2009.*
- White, M. (2007). *Making Search Work: Implementing Web, Intranet and Enterprise Search.* Facet Publishing.
- White, R. W., Dumais, S. T., and Teevan, J. (2009). Characterizing the influence of domain expertise on web search behavior. *Proceedings of the Second ACM Conference on Web Search and Data Mining (WSDM '09)*, pp. 132-141.
- White, R. W., Jose, J. M., and Ruthven, I. (2003). A task-oriented study on the influencing effects of query-biased summarisation in web searching. *Information Processing and Management*, 39, 5, pp. 707-733.
- White, R. W., and Morris, D. (2007). Investigating the querying and browsing behavior of advanced search engine users. *Proceedings of the 30th ACM SIGIR conference on Research and development in information retrieval*, pp. 255-262.
- White, R. W., and Roth, R. A. (2009). *Exploratory Search: Beyond the Query-Response Paradigm.* Morgan and Claypool.
- Wildemuth, B. M. (2004) The effects of domain knowledge on search tactic formulation. *Journal of the American Society For Information Science and Technology*, 55, 3, pp. 246-258.
- Xiang, B., Jiang, D., Pei, J., Sun, X., Chen, E., and Li, H. (2010). Context-aware ranking in web search. *Proceedings of the 33rd international ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '10*, pp. 451-458.
- Xu, J., and Croft, W. B. (1996). Query expansion using local and global document analysis. *Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '96*, pp. 4-11.

- Xu, S., Jin, T., and Lau, F. C. (2009). A new visual search interface for web browsing. *Proceedings of the Second ACM international Conference on Web Search and Data Mining*, pp. 152-161.
- Yee, K.P, Swearingen, K., Li, K. and Hearst, M. (2003). Faceted metadata for image search and browsing. *In Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '03)*, pp. 401-408.
- Zhang, L., and Zhang, Y. (2010). Interactive retrieval based on faceted feedback. *Proceedings of the 33rd international ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '10*, pp. 363-370.

APPENDIX A. SQL Experiment (see section 3.5)

A.1. Task Questions

Task 1: What is a view in relational databases? Define a view for the aircraft table, which contains the columns model and aircraft_name.

Task 2: Give the insert command to insert a new aircraft where the call_sign is Charlie-Tango, the model is Airbus330 and the aircraft_name is Killian into the table Aircraft.

Task 3: Create a table aircraft_costs containing attributes aircraft_type (Max 20 characters long), aircraft_purchase_cost (Max 20 characters long), aircraft_vendor (Max 20 characters long), aircraft_price (Integer), date_of_purchase (Date).

Task 4: Insert a new row into the Aircraft table containing values where the call_sign is BB-GGG, the aircraft_name is Blue Bird, the model is Boeing 747, the no_club_seats is 60, and the no_economy_seats is 350. What happens if you insert the following values in a row: call_sign is BB-EEE, the aircraft_name is Blue Bird, the model is Boeing 747, the no_club_seats is 60, and the no_economy_seats is 350?

Task 5: Give the SQL command to delete the Aircraft table. Are there any conditions under which the deletion of a table would not be allowed?

Task 6: What is a trigger? Explain how it can be used for automatically insuring integrity in a relational database. Give an example of a trigger command and explain how that example works.

A.2. System Usability Scale (SUS) (Brooke, 1996)

	Strongly disagree				Strongly agree
1. I think that I would like to use this system frequently	1	2	3	4	5
2. I found the system unnecessarily complex	1	2	3	4	5
3. I thought the system was easy to use	1	2	3	4	5
4. I think that I would need the support of a technical person to be able to use this system	1	2	3	4	5
5. I found the various functions in this system were well integrated	1	2	3	4	5
6. I thought there was too much inconsistency in this system	1	2	3	4	5
7. I would imagine that most people would learn to use this system very quickly	1	2	3	4	5
8. I found the system very cumbersome to use	1	2	3	4	5
9. I felt very confident using the system	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with this system	1	2	3	4	5

A.3. Additional Questions

Strongly
disagree

Strongly
agree

1. I found the search system returned relevant search results for my query.

1	2	3	4	5

2. I found the search system returned irrelevant search results for my query.

1	2	3	4	5

3. I found the presentation of the search results helpful.

1	2	3	4	5

4. What did you like most about the search system?

5. What did you like least about the search system?

6. Any additional comments?

APPENDIX B. Symantec Experiment 1 (see section 4.5)

B.1. Task Questions

Task 1

1) You just purchased Norton 360 and would like to customize the backup feature. For this you gather relevant information on the available content provided by Symantec. Summarise general information on the main functions of the backup feature (e.g. functionality, backup media, file types, etc.) of Norton 360.

2) After using the backup feature you decide to delete your old backup called 'Henry' and make a completely new backup.

If you feel you have not gathered enough information on how to solve this task please gather more information. Please indicate how to create and delete backups.

3) Due to limited space on your hard drive you start to consider online backup. You would like to find out how this works.

Gather general information on how Norton Online Backup is activated.

4) After investigating Norton Online Backup you find out that you have to use your Online Norton Account to activate Online Norton Backup.

What does the Online Norton Account keep track of?

5) You are interested in details about Online Norton Account and the experiences users have made. (For this you investigate the User Forum content.)

Identify at least two typical problems users have with Norton Online Backup.

6) Is it possible to delete files from the Norton Online Backup?

Task 2

1) You are interested in the Antivirus definitions of Norton 360. For this you gather relevant information. Summarise general information on Antivirus definitions (e.g. Virus definition updates, Antivirus protection settings etc.).

- 2) What feature is responsible for updating definitions?
- 3) Can you provide an introduction to the functionality of this feature?
- 4) What updates are downloaded by this feature in Norton 360?
- 5) Where can you find the information on what date Norton 360 has updated the virus definitions the last time?
- 6) You get the following error: “Antivirus Definitions not updating”.

Can you find possible causes and/or solutions to the error message (up to two)?

Task 3

1) You are interested in the Identity protection feature of Norton 360. You are not sure which one of the following activities is addressed by this feature.

- a. Safeguards against online identity theft
- b. Scans the incoming and the outgoing emails for malicious content
- c. Blocks the fake Web sites and Crimeware
- d. Stores and encrypts your passwords to prevent accidental disclosure to unknown sites or unauthorized sites

2) You are also interested in Norton Safe Web and would like to be able to answer which of the following features are provide by Norton Safe Web:

- a. Offers a Web site rating service that extends protection to everyday online activities: search, browse, transact, and interact
- b. Provides a trusted visual indicator on search results pages of search engines
- c. Protects from the sites that can infect the computers and misuse user’s personal information
- d. Protects the host file from any changes

3) Finally you are interested in the Portable profile feature. Which one of the following statements are true:

- a. You can use the same portable Identity Safe profile with different user accounts on the same computer as well as on different computers where a compatible Identity Safe component exists in a Norton product
 - b. Any device that appears as a “removable disk” in Windows, with at least 5MB free space, and has a drive letter can be supported
 - c. The support includes USB flash drives, portable USB hard drives, cameras, MP3 players, cell phones, Zip drives
 - d. Optical drives and a fixed hard drive are also supported if user has rights to write on them
- 4) Can you find a possible solution to this problem: 'How do I get my Norton Toolbar back?'
- 5) Is it possible to delete the Norton Toolbar?

Task 4

1) You have been using Norton 360 for some time but you would like to gather more information on how to do task scheduling.

You would like to know what the difference between idle time automatic tasks and scheduling tasks is? Please state the difference.

2) After investigating the feature you are more interested in specific settings. Please investigate the following activities:

- a) How do you schedule automatic Backups?
- b) How do you specify the idle time duration for automatic tasks?
- c) You wish to disable idle time scans?

3) Before disabling the idle time scans you are interested in its implications such as performance and general user experience with this feature.

Investigate the available content and indicate your opinion on disabling the idle time scans.

B.3. General User Satisfaction, Reaction & Comments

	Strongly disagree				Strongly agree
1. I had to search a lot before I found interesting content.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
2. I did well on the different tasks.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
3. Overall, I am satisfied with the system performance, assistance and guidance.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
4. The system guided me towards more personally relevant content.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
5. I found the interaction with the system motivating.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
6. I found the interaction with the system engaging.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
7. I found the interaction with the system fun.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
8. What features/characteristics did you like most about the system?					
9. What features/characteristics did you like least about the system?					
10. Any additional comments?					

B.4. Comparative Questionnaire 1

	Strongly disagree				Strongly agree
1. I found search system A returned relevant content more prominently.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
2. I found the presentation of the search results of system A more helpful.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
3. I found the composition/grouping of the search results of system A more accurate.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
4. I found the composition/grouping of the search results of system A more helpful.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
5. The answer structure and content of system A was matching my expectations more accurately.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
6. The answer structure and content of system A was matching my knowledge state more precisely.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
7. The answer structure and content of system A was more helpful in solving the tasks.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
8. The content composition generated by system A was easier to navigate.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
9. I felt more guided across the different content sources by system A.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	

B.5. Comparative Questionnaire 2

	Strongly disagree				Strongly agree
1. I had to search more with system A before I found interesting content.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
2. I did better on the different tasks with system A.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
3. Overall, I am more satisfied with the system performance, assistance and guidance of system A.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
4. System A guided me more towards personally relevant content.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
5. I found the interaction with system A more motivating.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
6. I found the interaction with system A more engaging.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
7. I found the interaction with system A more fun.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	
8. Any additional comments?					

APPENDIX C. General Support Questions (see section 5.2)

<p>1. Where would you expect to find relevant introductory/overview information for product features?</p> <p>Please order your choices from 1 (most likely) to 4 (least likely).</p>	<p>Product Manual/ Built-in help</p>	<p>Support Articles</p>	<p>Forums</p>	<p>Other Web Resources (please specify)</p>
<p>2. Where would you expect to find relevant instructions/how-to information for product features?</p> <p>Please order your choices from 1 (most likely) to 4 (least likely).</p>	<p>Product Manual/ Built-in help</p>	<p>Support Articles</p>	<p>Forums</p>	<p>Other Web Resources (please specify)</p>
<p>3. Where would you expect to find relevant problem solutions for product features?</p> <p>Please order your choices from 1 (most likely) to 4 (least likely).</p>	<p>Product Manual/ Built-in help</p>	<p>Support Articles</p>	<p>Forums</p>	<p>Other Web Resources (please specify)</p>

APPENDIX D. Product reaction cards (see section 5.3)

The complete set of 118 Product Reaction Cards				
Accessible	Creative	Fast	Meaningful	Slow
Advanced	Customizable	Flexible	Motivating	Sophisticated
Annoying	Cutting edge	Fragile	Not Secure	Stable
Appealing	Dated	Fresh	Not Valuable	Sterile
Approachable	Desirable	Friendly	Novel	Stimulating
Attractive	Difficult	Frustrating	Old	Straight Forward
Boring	Disconnected	Fun	Optimistic	Stressful
Business-like	Disruptive	Gets in the way	Ordinary	Time-consuming
Busy	Distracting	Hard to Use	Organized	Time-Saving
Calm	Dull	Helpful	Overbearing	Too Technical
Clean	Easy to use	High quality	Overwhelming	Trustworthy
Clear	Effective	Impersonal	Patronizing	Unapproachable
Collaborative	Efficient	Impressive	Personal	Unattractive
Comfortable	Effortless	Incomprehensible	Poor quality	Uncontrollable
Compatible	Empowering	Inconsistent	Powerful	Unconventional
Compelling	Energetic	Ineffective	Predictable	Understandable
Complex	Engaging	Innovative	Professional	Undesirable
Comprehensive	Entertaining	Inspiring	Relevant	Unpredictable
Confident	Enthusiastic	Integrated	Reliable	Unrefined
Confusing	Essential	Intimidating	Responsive	Usable
Connected	Exceptional	Intuitive	Rigid	Useful
Consistent	Exciting	Inviting	Satisfying	Valuable
Controllable	Expected	Irrelevant	Secure	

Developed by and © 2002 Microsoft Corporation. All rights reserved.

APPENDIX E. Symantec Experiment 2 (see section 5.5)

E.1. Task questions

Task 1

- 1) What feature do you use if you want to turn off alerts during important tasks?
- 2) What different types of modes are there for this feature?
- 3) You notice that the icon of Norton 360 has changed when this feature is enabled, is that supposed to happen?
- 4) You can disable silent mode from the Settings user interface. (Yes/No)
- 5) What does Quiet mode do? When does it turn on? Can you add programs for which you want Norton 360 to turn on Quiet Mode?

Task 2

- 1) How do you restore files from a dvd?
- 2) What is the difference between restoring files and using Autorun Restore?
- 3) What information is stored by Identity Safe?
- 4) How do you backup Identity Safe data?
- 5) How does this differ from exporting Identity Safe data?

Task 3

- 1) You get the following error message: LU1806. What feature has caused this problem? Can you find out how to solve this problem?
- 2) What is the difference between this general update feature and Pulse updates?
- 3) You are using your computer in a new location and get the following error: Unable to connect to network proxy server. What do you need to do to resolve this?
- 4) What is the automatic update frequency? Can you change this?

E.2. Pre-Questionnaire

1. How many of the following Norton 360 features do you know about?

Antivirus, Backup, Pulse Updates, Firewall, Insight Network, SONAR Protection

- a. 0 b. 1-3 c. 4+

2. How many of the following actions have you performed using Norton 360?

Installation, Virus scan, Firewall configuration, Update download, Data restoration, Backup configuration

- a. 0 b. 1-3 c. 4+

3. Would you be able to advise people on using and configuring Norton 360?

- a. yes b. no

4. How often do you use a computer?

- a. every week b. every day c. many times a day

5. How often do you browse the web?

- a. every week b. every day c. many times a day

6. How often do you use search engines?

- a. every week b. every day c. many times a day

7. Have you ever used advanced search engine features (e.g. using the '-' sign to specify unwanted terms)?

- a. yes b. no

8. What do you tend to do when encountering a software problem?

- a. self-help (through manuals, forums, web searches, etc.)
b. contact the help/call centre

9. Which of the following statements applies to you most?

- a. I like getting a quick how-to/fix without additional explanations.
b. I like understanding the cause of a problem that has occurred.

10. What is more important to you?

- a. A webpage lays out the content in clear sequential steps.
b. A webpage gives me an overall picture and relates the content to other subjects.

11. How would you generally understand new software features?

- a. Once I understand all the parts, I understand the whole thing.
b. Once I understand the whole thing, I see how the parts fit.

E.3. Interface Preference Prequestionnaire (asked for each of the screens)

Before starting the Norton 360 tasks, we are interested in your reaction to 4 different screens. Suppose you were looking for information about making ice cream.

What would your reaction be towards the following screen?

	Strongly disagree				Strongly agree
1. I know where to find relevant information to the given task.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5
2. I feel overwhelmed by the amount of information on the screen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5
3. I recognise what kind of information is linked from each part of this interface.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5
4. There is a lot of irrelevant information on the screen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5
5. I would prefer more guidance across the different information types and sources on the screen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5
6. Overall, I like this screen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5
7. Overall, I think this system would be easy to use.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5

Screen 1 (wdyl.com)

The screenshot shows a web browser window with the URL www.wdyl.com/#make+ice+cream. The page title is "What do you love? make ice cream". The search results are organized into several columns:

- Blogs with Blog Search:** Includes articles like "Big Dipper Ice Cream | Ice Cream Cone... I discovered Big Dipper during a visit to Missoula a few years ago. It's a...", "(Guest Post) Queen of Ice Cream Aa 32... It is no secret that Chloe is the queen of ice cream and frozen yogurt so wha...", "Simple Homemade Ice Cream | Make Home... Making Homemade Ice Cream can be a real treat for the whole family.", and "Fron Ice Cream | Holiday Online Busin... Prepare the ice cream nicely prematurely so as to make it consistent. It's...".
- News with News:** Includes "Ice cream now on menu at Burger King MiamiHerald.com - 44 minutes ago BY ELAINE WALKER Burger King wants to make sure it has something on its menu to appeal to everyone and what better way than ice cream. The Miami fa...", "The ice cream man cometh Dallas Voice - 23 hours ago Modesty aside, Duke's skills with frozen treats keep him busy, traveling the country and teaching restaurateurs and chefs how to make gelato, ice c...", and "Business briefs: Cold Stone Creamery hosts ice cream social Chicago Sun-Times - 1 day ago Cold Stone Creamery in downtown Naperville and the Make-A-Wish Foundation invite the community to help make wishes come true at the 10th annual Wor...".
- Books with Books:** Includes "How to Make Ice Cream: An Illustrated Step-By-Step Guide to Perfect Ice ... Cook's Illustrated Magazine, John Burgoyne, Christopher Kimball - 1997", "Ben & Jerry's homemade ice cream & dessert book Ben R. Cohen, Jerry Greenfield, Nancy J. Stevens, Lyn Severance - 1987", and "Ice Cream Mix-Ins: Easy Homemade Treats Jeff Keys, Zac Williams - 2009".
- Web results with Web Search:** Includes "Simple Homemade Ice Cream | Make Home... Making Homemade Ice Cream can be a real treat for the whole family.", and "Fron Ice Cream | Holiday Online Busin... Prepare the ice cream nicely prematurely so as to make it consistent. It's...".
- Discussions with Groups:** No results are visible.
- Videos on Youtube:** No results are visible.

Screen 2 (yippy.com)

The screenshot shows a web browser window with the URL search.yippy.com/search?input-form=clusty-simple&v%3Asources=webplus&v%3Aproject=clusty&query=make+ice+cream. The page title is "make ice cream". The search results are organized into a sidebar and a main content area:

- Sidebar:** Lists "All Results (174)" with categories: "Homemade Ice Cream (31)", "Ice Cream Recipes (20)", "Milk (13)", "Sale, Drinking (13)", "Ice Cream Parlour (16)", "Machines, Soft Serve Ice Cream (8)", "Summer (7)", "Vanilla Ice Cream (8)", "Frozen Yogurt And Ice Cream (8)", "Cone (8)", and "more | all clouds".
- Main Content Area:** Shows "Top 170 results retrieved for the query make ice cream (details)". The results include:
 - "Homemade Ice Cream in a Bag: Summer Crafts for Kids: Easy ... How to make it: Fill the large bag half full of ice, and add the rock salt. Seal the bag. Put milk, vanilla, and sugar into the small bag, and seal it. crafts.kaboose.com/ice-cream-in-a-bag.html - [cache] - Additional Sources, Gigablast"
 - "Making Ice Cream: Homemade Ice Cream - www.ice-cream-recipes.com Making ice cream - information on homemade ice cream, how to make, main ingredients, types of ice cream recipe, etc www.ice-cream-recipes.com/making_ice_cream.htm - [cache] - Additional Sources"
 - "Recipes for Making Ice Cream, Frozen Yogurt, and Sorbet Here is a collection of recipes we have compiled. If you don't find what you are looking for, send us an Email and we will do our best to add what you need when we do ... www.makeicecream.com/rectormakeice1.html - [cache] - Additional Sources"
 - "Make Ice Cream in a Baggie - Freezing Point Depression Make Ice Cream in a Baggie - Freezing Point Depression and Colligative Properties chemistry.about.com/cs/howto/a/aa020404a.htm - [cache] - Additional Sources"
 - "Homemade ice cream: how to make Learn how to make homemade ice cream from scratch. Tons of ice cream recipes. Your complete guide to making home made ice cream. Find reviews of ice cream ... make-ice-cream.com/default.aspx - Cached page make-ice-cream.com/default.aspx - [cache] - Additional Sources, Bing"
 - "Ice Cream Making Tip The other day I was making ice cream and I filled the frozen container to the very top. About 20 minutes later when I went to check on the consistency, I was shocked ... www.yumsugar.com/Ice-Cream-Making-Tip-8424968 - [cache] - Additional Sources"
 - "MakeIceCream.com - ice cream makers , ice cream maker ... Cuisinart ICE-50BC Supreme Commercial Quality Ice Cream Maker CUISINART ICE-50BC www.makeicecream.com - [cache] - Additional Sources"

Screen 3 (HowStuffWorks (1))

The screenshot shows a web browser window with the URL `www.howstuffworks.com/search.php?terms=make+ice+cream`. The page header includes the HowStuffWorks logo and a search bar. The main content area is titled "howstuffworks search results for: make ice cream" and indicates that 243 articles were returned. A sidebar on the left lists navigation options: Articles, Images, Videos, Expert Reviews, and Web Results. The main content area features several article snippets:

- How Ice Cream Works: Making Ice Cream**: Making ice cream commercially is actually quite similar to the process of making ice cream at home. Learn the steps of making ice cream... of making ice cream is basically the same. The only difference is the scale of the operation. First, you need ice cream mix. You can buy commercially... [Science > Edible Innovations](#)
- Introduction to How Ice Cream Works**: Ice cream is made of molecules of fat suspended in a structure of water, sugar and ice. Learn about the history of ice cream and see how ice cream...rozen dairy product [ref]. Although ice cream can be easy to make at home, it is actually a very complex substance. In this article, we'll learn how it's made... [Science > Edible Innovations](#)
- How Ice Cream Works: Five-minute Ice Cream**: Five-minute ice cream can be made at home without an ice cream freezer. Learn how to make delicious five-minute ice cream with a bag and everyday... There are many recipes out there for making your own ice cream at home, but did you know that you can make your own ice cream in five minutes... [Science > Edible Innovations](#)
- How Ice Cream Works: Ice Cream Business and History**: The history of ice cream began in the 1700s although people have long enjoyed flavored ices. Learn about the history of ice cream and the modern ice...of the ice cream cone is controversial. An Italian immigrant named Italo Marchionni has a strong claim because he filed a patent for a cone making machine... [Science > Edible Innovations](#)
- How to Remove Ice Cream Stains: More Tips on Removing Ice Cream Stains**: Ice cream stains can be a subtle hint to avoid sweets, but they can be cleaned from any material or surface. Learn how to remove ice cream stains... as soon as possible. When you scream for ice cream, make sure it's not because it's dripping all over your clothes. But if you are that unlucky, use... [Home & Garden > Food Stains](#)

Screen 4 (HowStuffWorks (2))

The screenshot shows the full article page for "How Ice Cream Works" by Ed Grabianowski. The URL is `science.howstuffworks.com/innovation/edible-innovations/ice-cream.htm`. The page includes a breadcrumb trail: Home > Science > Innovation > Edible Innovations. Below the title, there are social sharing options for Print, Cite, Facebook, Share (4), and Tweet (0). A "What do you know about salt?" quiz is featured on the right side. The article text begins with: "The U.S. ice cream industry sells more than a billion gallons of ice cream each year, dispensing cones, gallons, pints, sundaes and other desserts through grocery stores and ice cream shops. In fact, eight percent of all the milk produced in the United States ends up in a frozen dairy product [ref]."

Although ice cream can be easy to make at home, it is actually a very complex substance in this article, we'll learn how it's made, what goes into it and who invented it. We'll also learn how to quickly make ice cream in your kitchen.

Ice Cream or Frozen Dessert?

Not just any frozen treat can be called ice cream. In fact, the [U.S. Department of Agriculture](#) has specific rules that define what can and can't be labeled "ice cream." To bear the "Meets USDA Ingredient Standard for Ice Cream" stamp, it has to contain at least 10 percent milk [fat](#), and a minimum of six percent non-fat milk solids. A gallon has to weigh at least 4.5 pounds.

The range of milk fat (sometimes referred to as butter fat) used in ice cream can go from the minimum 10 percent to a maximum of about 16 percent. Most premium ice creams use 14 percent milk fat. Higher fat content leads to better, richer taste and a creamier texture. Ice cream makers don't go higher than 16 percent because it would be costly and very high in [calories](#). An ice cream with this much milk fat would also taste so rich that people would probably eat it in smaller amounts, which would be bad news for people who sell ice cream for a living.

Other frozen desserts, such as sorbets, low-fat ice cream, and frozen yogurt, are not technically

Ice Cream Oddities

©2009 HowStuffWorks
Ice cream is a welcome summer treat. See more pictures of classic snacks.

E.4. Post-task Questionnaire

Strongly
disagree

Strongly
agree

1. The task was complex.

1	2	3	4	5

2. I did well on the task.

1	2	3	4	5

3. I found the result pages generated for me helpful in solving the task.

1	2	3	4	5

4. The result pages guided me towards content that was relevant to the task.

1	2	3	4	5

E.5. Application-specific Questionnaire

	Strongly disagree				Strongly agree
1. The generated result pages were easy to navigate.	1	2	3	4	5
2. I found the amount and diversity of content overwhelming on every screen.	1	2	3	4	5
3. I found content that was relevant to my task easily.	1	2	3	4	5
4. I found the way the result pages were composed and presented unclear and inconsistent.	1	2	3	4	5
5. I found the generated result pages to be clean.	1	2	3	4	5
6. I had to search a lot before I found relevant content.	1	2	3	4	5
7. I found the search system returned relevant content more often than irrelevant content.	1	2	3	4	5
8. I often used the query intent option (i.e. "I want to...") to narrow down the search results.	1	2	3	4	5
9. The system generated appropriate presentations for the chosen query intents.	1	2	3	4	5
10. I felt guided across the different content sources.	1	2	3	4	5
11. Overall, I found the interaction with the system frustrating.	1	2	3	4	5
12. Overall, I found the interaction with the system motivating.	1	2	3	4	5
13. What features/characteristics did you like most about the system?					
14. What features/characteristics did you like least about the system?					
15. Any additional comment					

APPENDIX F. Full Pre-questionnaire results (see section 5.5.3)

Would you be able to advise people on using and configuring Norton 360?*			
	C1	C2	C3
Yes	9	7	8
No	21	19	23
Have you ever used advanced search engine features (e.g. using the '-' sign to specify unwanted terms)?**			
	C1	C2	C3
Yes	20	19	21
No	10	7	10
What do you tend to do when encountering a software problem?			
	C1	C2	C3
Self-help (through manuals, forums, web searches, etc.)	29	25	27
Contact the help/call centre	1	1	4
Which of the following statements applies to you most?			
	C1	C2	C3
I like getting a quick how-to/fix without additional explanations.	16	15	16
I like understanding the cause of a problem that has occurred.	14	11	15
What is more important to you?			
	C1	C2	C3
A webpage lays out the content in clear sequential steps.	25	20	27
A webpage gives me an overall picture and relates the content to other subjects.	5	6	4
How would you generally understand new software features?			
	C1	C2	C3
Once I understand all the parts, I understand the whole thing.	17	18	19
Once I understand the whole thing, I see how the parts fit.	13	8	12

APPENDIX G. Symantec Experiment 3 (see section 6.5)

G.1. Task questions

Task 1

- 1) Wie kann man Identity Safe Daten sichern?
- 2) Was ist der Unterschied zwischen dieser Funktion und dem Exportieren von Identity Safe Daten?
- 3) Gibt es einen Phishing Schutz den man kostenlos testen kann?
- 4) Was sind die Haupteigenschaften vom Phishing Schutz?
- 5) Wie bekommt man die Norton Symbolleiste zurück? Wie haben andere Leute dieses Problem behoben?

Task 2

- 1) Welche Produktfunktion benutzt man um Norton 360 zu aktualisieren?
- 2) Was ist der Unterschied zwischen dieser Funktion und Pulse Update Aktualisierung?
- 3) Mit welcher Update Häufigkeit wird Norton 360 aktualisiert?
- 4) Kann man eine externe Festplatte optimieren?

Task 3

- 1) Welche automatischen Aufgaben kann man mit dem Produkt planen?
- 2) Wie kann man ein Backup wiederherstellen? Was ist der Unterschied zwischen einer Backup Wiederherstellung und einer Wiederherstellung mithilfe von Norton 360 Autorun Restore?
- 3) Sie haben alle Online Backups gelöscht. N360 zeigt jedoch an, dass Sie über keinen freien Speicherplatz verfügen. Wie haben andere Leute dieses Problem behoben?
- 4) Sie erhalten die folgende Fehlermeldung: 'Fehler 3043'. Wie kann man dieses Problem beheben?

G.2. Pre-Questionnaire

1. How many of the following Norton 360 features do you know about?

Antivirus, Backup, Pulse Updates, Firewall, Insight Network, SONAR Protection

- a. 0 b. 1-3 c. 4+

2. How many of the following actions have you performed using Norton 360?

Installation, Virus scan, Firewall configuration, Update download, Data restoration, Backup configuration

- a. 0 b. 1-3 c. 4+

3. Would you be able to advise people on using and configuring Norton 360?

- a. yes b. no

4. How often do you use a computer?

- a. every week b. every day c. many times a day

5. How often do you browse the web?

- a. every week b. every day c. many times a day

6. How often do you use search engines?

- a. every week b. every day c. many times a day

7. Have you ever used advanced search engine features (e.g. using the '-' sign to specify unwanted terms)?

- a. yes b. no

8. What do you tend to do when encountering a software problem?

- a. self-help (through manuals, forums, web searches, etc.)
b. contact the help/call centre

9. Which of the following statements applies to you most?

- a. I like getting a quick how-to/fix without additional explanations.
b. I like understanding the cause of a problem that has occurred.

10. What is more important to you?

- a. A webpage lays out the content in clear sequential steps.
b. A webpage gives me an overall picture and relates the content to other subjects.

11. How would you generally understand new software features?

- a. Once I understand all the parts, I understand the whole thing.
b. Once I understand the whole thing, I see how the parts fit.

12. What is your level of proficiency in English?

- a. BASIC: You can communicate in predictable contexts and on familiar topics, but with some difficulty.
- b. MODERATE: You can communicate comfortably in familiar social and work situations.
- c. HIGH: You can communicate effectively in most social and work situations.
- d. NATIVE

13. What is your level of proficiency in German?

- a. BASIC: You can communicate in predictable contexts and on familiar topics, but with some difficulty.
- b. MODERATE: You can communicate comfortably in familiar social and work situations.
- c. HIGH: You can communicate effectively in most social and work situations.
- d. NATIVE

G.3. Post-task Questionnaire

Strongly disagree

Strongly agree

1. The task was complex.

1	2	3	4	5

2. I did well on the task.

1	2	3	4	5

3. I found the result pages generated for me helpful in solving the task.

1	2	3	4	5

4. The result pages guided me towards content that was relevant to the task.

1	2	3	4	5

5. I found it easy to navigate across German and English content.

1	2	3	4	5

6. I would like the presentation of German and English content more clearly separated from each other.

1	2	3	4	5

7. I would like the German and English content to be more integrated.

1	2	3	4	5

8. I felt guided across the German and English content.

1	2	3	4	5

9. I found the German and English content to be highly related to each other.

1	2	3	4	5

10. Any additional comments?

G.4. Comparative Questionnaire

1. Overall, which system did you prefer the most?
 - a. System A
 - b. System B
 - c. System C
2. Overall, which system did you prefer the least?
 - a. System A
 - b. System B
 - c. System C
3. In which system was it easiest to navigate between German and English content?
 - a. System A
 - b. System B
 - c. System C
4. Which presentation provided the most guidance across German and English content?
 - a. System A
 - b. System B
 - c. System C
5. I felt overwhelmed when different languages were shown on the same screen.
6. I like it when the content is presented in the language it was originally created.
7. I would prefer to have all content presented in a single language.