Hamlet goes XML: CrossAnnotationLinking and Personal Learning Experiences

Benjamin Birkenhake[1], Stefanie Panke[2], Andres Witt[1]
Universität Bielefeld[1], Institut für Wissensmedien (IWM)[2]
benjamin.birkenhake@gmail.com, s.panke@iwm-kmrc.de, andreas.witt@uni-bielefeld.de

Abstract

XML-based technologies offer powerful resources for open source applications in the field of e-learning. The paper describes a model of hypertext as interlinked structures that can be intertwined by cross-annotation linking. This infrastructure integrates multiple perspectives and allows creating a personal learning environment. We exemplify the approach in a case study: the Hamlet project. In the course of this project, several German translations of William Shakespeare's Hamlet have been collected and annotated. Two different annotation layers are used to achieve a cross-linking reference between the various German translations. We will describe the theoretical background of cross-annotation linking and the actual technological implementation of the system. Additionally, we will use the personas method to gain insights into the potential benefit of the system as a personal learning environment.

1. Introduction

Although in theory open source and e-learning mix quite happily, the day-to-day practice often presents a different picture. Users have a more or less justified prejudice about open source being "somewhat difficult" and only for "techies". In this article we will reflect upon typical problems and propose a treatment for curing. We will describe the development of an XML-based environment from two different perspectives: The technological point of view and the user-centred design perspective.

XML allows combining single source storage of data with multiple modes of presentation. This offers rich opportunities for e-learning with respect to personalization and information retrieval. Since locating information is a vital part of the reading process (Guthrie & Mosenthal, 1987), functions that allow for the retrieval and management of resources are core features of an e-learning environment. To reveal the full potential of XML-technology for these purposes, we had to shift our traditional view on hypertext as interlinked segments of text to a new metaphor of interlinked trees which is exemplified on the basis of the case study Hamlet (see section 2).

The digital technologies – especially mark-up languages – have changed the way information is managed and a growing body of literature has become available which tackles the technical, conceptual and pragmatic consequences of this development. Information retrieval focuses on structuring and presenting information in the digital environment and covers different aspects of organizing information like classification and taxonomy, indexing and thesaurus construction, as well as metadata (Marchionini, 2004). Information seeking models describe the processes of finding information to fill a certain knowledge gap from a user's point of view (see Marchionini, 1995; Ellis & Haugan, 1997). A wider perspective on information activities, including publishing and knowledge exchange, is given by theories on information behaviour (see e.g. Hektor, 2003). To reflect upon the information behaviour of the potential users of the hamlet corpus, we will deploy the personas method and match fictional users' needs with actual and desirable features of the Hamlet environment.

2. Hypertext as Interlinked Structures: The Case Study Hamlet

The way we search the web or other digital environments is contingent to the explicit or implicit concept of hypertext, which designers have in mind while developing digital artefacts.

Since more and more software is available to conveniently produce hypertext environments, hypertext has become a widely-used (text-)technology. At this point, new metaphors are needed for innovative features, especially for applications in the field of data mining and content management. In the following we will outline a perspective on hypertext as multi-rooted trees which are intertwined by cross-annotation linking (see Witt, 2005a; Witt 2005b).

2.1. Multi-Rooted Trees and Cross Annotation Linking

What is cross-annotation linking? This question is answered best within a historical flashback: The discussion of the first OHCO (Ordered Hierarchy of Content Objects) model (see De Rose et al., 1990; Renear et al., 1996) has shown that textual data could not be understood as a merely single hierarchy of content objects. This evidence was underlined by the emergence of new technologies. The dissemination of SGML and XML put forth an increasing application of multiple annotations of one source of textual data. These multiple annotations usually contain different layers of information, e.g. textual structure and linguistic, philological or narrative information. The different annotations form several trees above the primary textual data. The branches and leaves of these trees can be related to one another in several ways: One or more branches of one tree can be part of a branch of another tree. Two branches can have the same starting point and a different ending, as well as different starting positions and an identical ending - which means that branches can overlap (see Durand, 1999; Duruseau & O'Donnell, 2002). This typology of relations - together with the schemas of the different annotation layers - can be used to (semi-)automatically generate hyperlinks between the nodes of the different trees. This allows creating a network of multiple perspectives on one text being linked to one another. As a result, hypertext is no longer based on links between nodes, but offers a reference mechanism between perspectives.

2.2. The Case Study Hamlet

Although Shakespeare's Hamlet is obviously not a unique Hypertext, it is an interesting object to test cross-annotation-linking and several other hypertext-technologies. There is no original edition of Hamlet, which was authorized by Shakespeare during his lifetime. We only have different print editions, which all have a different status concerning their quality, overall length, content and storyline. The most important among these are the so called first folio, the first Quattro and the second Quattro edition of Hamlet. During the centuries editors tried to combine these early Editions to the best Edition possible. A comparable situation exists within in the field of German translations of the play. Almost every Translator used several of the early English editions as a basis for a new translation. This leads to a situation in which almost every German or English edition of Shakespeare's Hamlet is a composition of several sources. The relation the editions have with their sources and with each other form a wide network, which could be presented in an e-learning-environment.

Another interesting aspect of Shakespeare's Hamlet is the outstanding position the play has within the western culture for centuries. Hamlet is the single most researched piece of literature, has been analyzed in from various perspectives and is a part of western education. This leads to the request, that a digital environment should represent the variety of perspectives on the play. As part of a PhD thesis, the most important English editions and several German translations of William Shakespeare's Hamlet have been collected and annotated in different ways. This leads to a corpus of Hamlet editions in which each text may exist in multiple forms.

Basis for the XML-annotations are text files, which are transformed to XML using regular expressions. The basic XML-format is TEI 4 drama base tag set. TEI 4 is a major open source concept of the Text Encoding Initiative. It contains a modular Schema for several document-classes. The drama base tag set offers almost all tags needed for a general, formal annotation of a play. In order to provide an easy to annotate mechanism to represent the translation- or origin-relation between lines or paragraph within editions on the one hand and the sources on the other hand, some attributes were added to TEI by us.

The TEI-annotated documents are used for further annotations and presentation. The TEI-documents were automatically enriched with further mark-up, using an open source "auto-tagger". This auto-tagger annotates single words, including the part of speech and the principle form. The TEI-documents are also the basis for the XHTML-presentation. As the TEI-structure contains all information necessary for a graphical presentation, these documents are transformed to XHTML, which is used to present the corpus. This transformation is made with

several XSLT-Style sheets. In the same way XSL-FO is used to generate PDF-versions of each edition. Table 1 illustrates which annotations are provided for the different editions.

Edition	Txt	TEI	XHTML	STTS	Narration
1st Folio	✔	✔	✔		
1st Quattro	✔	✔	✔		
2nd Quattro	✔	✔	✔		
Moby	✔	✔	✔		
Wieland	✔	✔	✔		✔
Schlegel	✔	✔	✔	✔	✔
Fontane	✔	✔	✔	✔	✔
Hauptmann	✔	✔	✔	✔	

Table 1: annotation layers for the different editions.

In many cases translators have re-arranged the flow of stanzas or the course of action. Therefore it is useful to provide an alternative linking mechanism, which does not only focus on the language and the formal structure, but also on the plot. To provide this reference the narrative information is annotated in another layer. This allows to find the same event in different translations of the play. The narrative annotation layer basically consists of events, which can be seen as the smallest elements of the plot. Several events form an action. These can be grouped into more complex actions. All elements are embraced by the root element translation.

Obviously, events may start within one line and end several lines or even speeches later. Since the narrative structure is overlapping with the TEI, both could be stored in separate annotations. Scenes can provide a meaningful unit for basic parts of the plot. Thus the formal and the narrative annotation are semantically aligned - in addition to their reference on identical textual data. This relation can be exploited by creating links between the concept of a scene and the concept of specific actions. The respective linking mechanism is located on a Meta level: it operates on the schemas themselves and not on their instances. The references are generated mechanically on the Meta level, linking different perspectives together. Readers can explore the relations between events and scenes. The procedure could also be used to create a recommendation system as e.g. proposed by Macedo et al. (2003): the annotation integrates the knowledge of experts on narrative structures in the play Hamlet and provides this information to the reader. Figure one shows an example for two different annotation layers which are overlapping.

```xml
<event>
    <event unit="u0024" sequel="u0030" >HORATIO: Das kann ich; wenigstens das Gerede geht so.</event>
    <event unit="u0013" sequel="u0015">Unser letzter König, dessen Bild soeben vor uns erschien,</event>
    <event unit="u0025" sequel="u0026">ward-wie ihr wißt -vom Norweg Fortinbras - durch wett-
eifernden Stolz dazu gespornt - zum Kampf gefordert, </event><event unit="u0026" sequel="u0027">in
welchem unser tapfrer Hamlet (nach unsrem Wissen
schätzte ihn die Welt von dieser Seite) den Fortinbras
schlug; </event><event unit="u0025" sequel="u0026"> welcher - nach einem festgestellten Vertrage - [...] </event>
</event>


<sp n="48" f1nr="" q1nr="" q2nr="" moby="48" who="">
    <speaker>HORATIO</speaker>
    <l n="91" f1nr="" q1nr="" q2nr="" moby=""> Das kann ich; wenigstens das Gerede geht so.</l>
    <l n="92" f1nr="" q1nr="" q2nr="" moby="">Unser letzter König, dessen Bild soeben vor uns erschien,</l>
    <l n="93" f1nr="" q1nr="" q2nr="" moby="">ward-wie ihr wißt -vom Norweg Fortinbras - durch wett-</l>
    <l n="94" f1nr="" q1nr="" q2nr="" moby="">eifernden Stolz dazu gespornt - zum Kampf gefordert, in</l>
    <l n="95" f1nr="" q1nr="" q2nr="" moby="">welchem unser tapfrer Hamlet (nach unsrem Wissen</l>
    <l n="96" f1nr="" q1nr="" q2nr="" moby="">schätzte ihn die Welt von dieser Seite) den Fortinbras</l>
    <l n="97" f1nr="" q1nr="" q2nr="" moby="">schlug; welcher - nach einem festgestellten Vertrage -</l>
    [...]
</sp>
```

Figure 1: A multi rooted tree above a single textual data

2.3. Information Retrieval within the Hamlet Environment

As a first result of the multiple annotations, we got a corpus which is based on XML-technology and available via the web. As a second result we developed methods to cope with multiple annotated documents – a task which will probably be necessary more frequently with the growing popularity of XML-technologies. Especially the integration of the narration-annotation-layer can be seen as an example for further parallel annotations. The methods described above lead to an environment which offers different types of users different perspectives on a single, textual object. The following options result from the annotation:

1. The common TEI-annotation allows a structural linking-mechanism between the editions. This allows a user to jump from the first scene in the second act of one edition to the same scene in another edition.
2. Alternatively this annotation can be used to present the user a part of the play in on or more editions of his choice. For example he could choose the second scene of the second act and the editions »moby«, »1st folio«, »schlegel« and »fontante«. So different pieces of text would be presented as parallel columns. This interactive environment is created with PHP5 using the new "fast and simple" XML Extension.
3. The narrative annotation-layer allows several ways to explore a single text or compare some texts with each other. In the first case, the annotation of events and actions provides a way of comparing different editions esp. translations. It allows the user to jump for one edition to the same event in another edition, no matter in which part of the play the event is placed or of which words it is made of.
4. Using SVG an XML-based format for graphics the narrative structure of each translation could be visualized, ignoring the textual basis. This gives an »overview« of plot of the current edition.
5. The introduced concept of cross annotation linking allows us to offer the user automatically generated links from one annotation to another.

With this set of different linking-concepts we present users with almost complete freedom to explore the corpus in a way that fits to their needs. Every layer of information offers a way to access information of another layer in a different perspective. The resulting hypertext structure can be describes as a rhizome. This organization type is characterized by the multi-selective options resulting from a decentralized network of nodes and links. Everything can be linked with everything. In the epistemological philosophy of Gilles Deleuzes and Félix Guattari the rhizom forms the counterpart of the metaphor of the knowledge tree. The latter symbolizes formalized, structured, subject-oriented thinking, whereas the rhizom is a symbol of nomadic thinking (Berressem, 2000).

3. Designing for Learning: The Personas Method

We have seen that cross-annotation linking can provide an infrastructure for personalized learning environments. Nevertheless, the application of high-end technologies can result in very poor learning settings, because the development process is focussed less on the users and their goals and more on the computational challenges of an ongoing project. How can this narrow perspective be avoided or ameliorated? This section describes the personas approach as a method of user-centred design. Personas are fictional users who have individual goals and needs which are reflected in their usage of an environment (Cooper, 1999). They avoid designing for oneself and help the developer to decide which functions will be fruitful for the persons who have to work or learn with the technological artefacts (s) he produces.

The personas technique is an established method in product design (Sinha 2003) and "a powerful complement to other usability methods" (Pruitt & Grudin 2002). The application of personas - fictional people – to represent an abstract consumer has a long tradition in the field of marketing (Pruitt & Grudin 2002). At the end of the past century the use of personas was also heralded in product design (Sinha 2003). Meanwhile the method is applied in different contexts, for example in technical writing to create user documentation (Calde 2004) and is used as well as a design method in software engineering.

What is the advantage of using personas as compared to identifying target groups and designing for these different user groups? The personas approach tries to engage and immerse the designer in the everyday-life of potential users. "Personas help define the product by replacing the abstract, elastic user with the vibrant presence of a specific user who becomes part of the design process" (Sinha, 2003). As a kind of projection foil, personas as elaborated archetypes serve to identify the (information) needs and possible behavioral patterns of the potential users. "Understanding user information needs and mental models is important for design in information-rich domains" (Sinha 2003). Functionalities may be derived easier following the personas needs, interests and possible actions than in abstract design processes.

3.1. Personas for the Hamlet Environment

In order to gain reliable results for the design process, personas cannot be built from scratch – they have to be generated based on data. At this, qualitative data is suitable for the modelling of the personas' characters and quantitative data may be used – if necessary – for the weighting of the different personas developed (see Arnold et al., 2005). Since in the case study no explicit target group analyses were accomplished, the personas were created on the basis of experience reports: In the context of the Hamlet project several seminars were held within the period from winter semester 2003 to summer semester 2004. The courses on "Cyberhamlet" (http://www.cyberhamlet.net/) were concerned with the structure of the annotated corpus as well as a graphic conversion of the drama into a comic. Participants were bachelor and masters students from a wide range of subjects, such as media design, text technology, literature, English Studies, etc. The lecturer was interviewed with respect to his experiences form the courses. According to the objectives and interests of the students who attended these seminars two personas were created. The goal was to model two rather disparate user-profiles, to see how the Hamlet environment would fit to the respective needs. In the following you will meet Beatrice and Gerd, two fictional users with distinctive backgrounds and expectations concerning the interface options provided by the Hamlet environment.
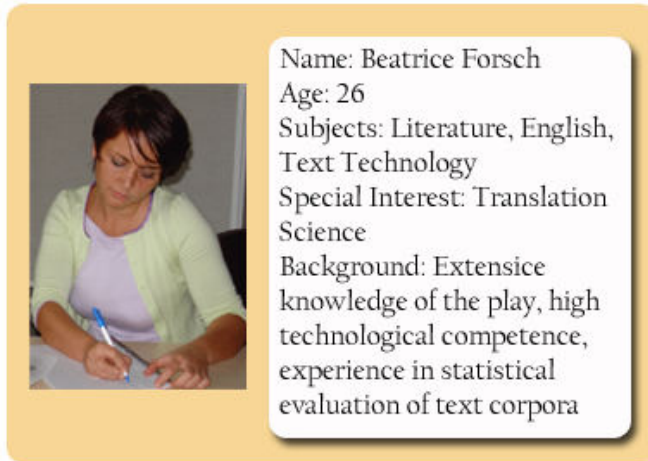
Fig. 2: Persona Beatrice Forsch

Beatrice Forsch, 26, studies comparative literature science, English and text technology at the University of Bielefeld with the goal to obtain a master degree. Her special interest is the translation science. She has advanced theoretical knowledge of English history and literature and has experiences in using computers for statistical analysis of textual data. In her master thesis she wants to analyze and show that each translation is at the same time an interpretation. The material offered within the Hamlet corpus is predestined as a case study, since she can work on the corpus with different statistical methods. Beatrice is especially interested in the scientific expertise coded into the narrative layer of annotation. She wants to compare the flow of events between translations. For further analysis on the differing translations, she would like to conduct semantic field analysis, which could be provided through additional text mining tools. For her purposes it is important that she can export the results of her work in different formats: She wants to print out graphic overviews of the sequence of events to get a quick overview on interesting differences between the translations. To include the material in her thesis she wants to export segments of the graphical maps as GIF, PNG and SVG files. Furthermore quotes from the corpora should be exportable as text files and contain the respective source information.

Fig.3: Persona Gerd Ravig

Gerd Ravig, 23, is a bachelor student of media design at the University Of Applied Science of Bielefeld. Gerd is very practice-oriented: From his point of view, scientific theories should be directly convertible into his creative working context. Together with fellow students Gerd produces filmlets, animation and trick films. He is interested in the material "Hamlet", since he has seen the movie "The Denmark Company", a modern adaptation of the play. Gerd works with the corpus in the context of a project seminar. As a project assignment, he works on a

Comic version of the Hamlet. In the context of the course Gerd would like to complement his practical knowledge in the production of films with theoretical know-how on the design of film script. On the basis the Hamlet environment he is to learn fundamental narrative concepts such as events, actions, chains of events and character constellations. Since Gerd likes to work with graphical representations he uses the possibility to visualize for example event segments from the data. To get familiar with narrative theories, Gerd needs additional didactical tools, e.g. a glossary should clarify the theoretical terms and additional information should be provided which helps him to re-contextualize the action of the play within its historical setting.

3.2. Information Seeking Behavior of the Personas

To identify useful features for an e-learning environment based on multi-rooted trees we apply an adapted version of Ellis' model of information seeking (Ellis & Haugan, 1997; Choo et al., 2000) to specify user tasks within the environment.

Process	Beatrice (Using the Hamlet Corpus as a Tool for Research and Explorative Learning)	Gerd (Using the Hamlet Corpus as an Instructional Learning Environment)
Starting	Refining the research questions she wants to analyze within the Hamlet environment, getting a first overview of the material by retrieving general statistical information on the translations (e.g. number of words, scenes, entries, etc.).	Getting an overview of the play's structure by using the graphical representation as an overview or index. Identifying personal points of interest through mixed modality representations, e.g. a movie version of Hamlet which is linked to the textual data.
Chaining	Following links between the various layers of annotation, identifying points of interest, using the visual representation of narrative structures to compare e.g. Schlegel's and Fontane's translation of the play	Following links between the various layers of annotation, especially exploring didactical context-information and glossary explanations.
Browsing	Investigating specific characters more closely, e.g. by retrieving a selection of all scenes of Rosenkranz and Güldenstern.	Following a certain chain of events, working through a pre-defined learning path which explains and illustrates the relationship between Hamlet and Ophelia.
Differentiating	Identifying and bookmarking specific segments, saving them in the personal profile. Using the data mining filter to scan for the occurrence of semantic fields.	Reading specific segments, express questions and ideas through annotations, saving reading paths.
Monitoring (Adding)	Reviewing previous results stored in the personal profile. Obtaining information about new features or the work of fellow students, learning how to program own queries in XSLT through embedded tutorials, providing own results within the environment.	Re-reading saved paths and personal comments stored in the personal profile. Obtaining information about new features or the work of fellow students, adding own graphical work, using export function to add speeches into the comic produced in the project assignment.
Extracting	Exporting results of analysis through screenshots, quotes and different export formats offered.	Selecting useful pages and sites by bookmarking, printing, copying and pasting, etc.

Table 2: Information Seeking Behaviour (adopted from Choo et al. 2000, 7) applied to E-Learning Tasks within the Hamlet Environment

3.3. Recommendations for Further Development

The Hamlet corpus is still work in progress. A very basic interface has been provided based on the XHTML presentation generated from the annotated corpus. So at a very early stage of development useful features for learning purposes are discussed. The following summary

comprises central recommendations for employing the Hamlet corpus in an e-learning context.

As we have learned from the scenarios of use sketched with the personas, personal interests, learning goals and information seeking motives may differ widely among the potential users of the system. Therefore it is of vital importance that the students can select which perspective on the corpus they want to pursue. They should be able to reduce the complexity of the interface according to their specific needs. Specific browsers should allow to select specific translations and perspectives. Browsing "perspectives" would allow students to switch between e.g. the TEI annotation layer and the narrative information. Browsing "translations" would provide the possibility to compare e.g. the narrative structure of Schlegel's and Fontane's translation. Moreover, visually coded links and a colour-navigation may help to reduce cognitive load and at the same time provide explorative flexibility.

Users like Gerd, who are not thoroughly familiar with the play, will actually want to read parts of the play within the learning environment. Therefore it is important that the web-interface provides reader-friendly typeset, as well as printable versions. To resume the reading process at a later date, students should be able to set bookmarks and annotate single nodes or segments. Furthermore, novice students would profit from pre-defined learning and reading paths, which illustrate a certain question comprehensively, e.g. the relationship between Hamlet and Ophelia. Additional information like glossary items explaining narrative concepts as well as information on the social and historical background of the play would support and deepen the understanding.

Advanced students like Beatrice will use the environment to investigate very specific questions. Hence, filters should be provided to extract meta-information and to assembly and display only specific segments of the textual data, e.g. all scenes of Rosenkranz and Güldenstern. A "data mining tool" could offer the possibility to define semantic fields and analyse different translations of the play accordingly. To give advanced users even more flexibility, tutorials could be produced which explain how to form a query with programming language XSLT, working directly on the annotated data.

4. Conclusions and Perspectives

We have presented a twofold strategy for the application of open source in the field of e-learning: We have described an XML application which reveals opportunities for information retrieval and personalization features. We assume that with growing possibilities to create multi-rooted trees on single-source primary data, the importance of cross-annotation linking will increase.

To provide recommendations for the use of cross annotation linking as an e-learning tool, we exerted the personas approach. Narrative design can serve as an instrument for the development of open-source based and user-friendly e-learning environments. The personas reflect the experience of working with students of the study course text technology at the University of Bielefeld. Further research should lead to a refined picture of the personas generated in this context. In this respect, involving other subjects through interviews with teachers and students may provide an interesting perspective.

To meet the recommendations derived from the personas approach still remains a challenge. Nevertheless, the evaluation of the features based on the personas helped the developers to understand the potential benefits and obstacles in deploying the Hamlet corpus as a self-directed learning environment. It is well-known that interface design is more than "lipstick on a bulldog". However, many technologically innovative projects suffer from weaknesses in providing an aesthetically attractive surface, which offers functionality really needed by its users.

In the case of the Hamlet environment, involving the future user in the design process took place before extensive work on the interface was accomplished. Therefore it is possible to balance between technological expenditures of implementation and didactical desirability. The proposed strategy will result in a feature specification which comprises technological, organizational and user-oriented perspectives.

5. References

Arnold, P., Geiser, B., Panke, S. (2005). Personas im Designprozess einer E-Teaching Community. In: J. M. Haake, U. Lucke & D. Tavangarian (eds.). *Proceedings DELFI 2005, 3. Deutsche e-Learning Fachtagung Informatik der Gesellschaft für Informatik e.V. (GI)*.

Berresem H. (2000). Unterwegs im Docuversum. Zur Topologie des Hypertext.. In: M. Klepper, R. Mayer & E.P. Schneck. (eds.). *Hyperkultur. Zur Fiktion des Computerzeitalters.* Berlin: de Gruyter.

Calde, Steve (2004) Using Personas to Create User Documentation. [Online] Available: http://www.cooper.com/content/insights/newsletters/2004_issue04/Using_personas_to_create_user _docs.asp [11/11/2004]

Choo , C.W., Detlor, B. and Turnbull, D. (2000). Information Seeking on the Web: An Integrated Model of Browsing and Searching, First Monday, 5/2, Online: http://firstmonday.org/issues/issue5_2/choo/index.html [Available, last checked 5/9/2005].

DeRose, S., Durand, D., Mylonas, E. and Renear, A. (1990). What is Text, Really? *Journal of Computing in Higher Education*, 1, 2, pp. 3-26.

Durand, D. G. (1999). *Palimpest: Change-Oriented Concurrency Control for the Support of Collaborative Applications*, PhD Thesis, Boston University.

Duruseau, P. and O'Donnell, B. M. (2002). Concurrent Markup for XML documents. In: *Proceedings of XML Europe 2002 Conference Proceedings* (Barcelona, Spain, May 20-23, 2002). Online: http://www.idealliance.org/papers/xmle02/dx_xmle02/papers/03-03-07/03-03-07.html [Available, last checked 5/9/2005].

Ellis, D. and Haugan, M. (1997). Modeling the Information Seeking Patterns of Engineers and Research Scientists in an Industrial Environment. *Journal of Documentation*, 53, pp. 384-403.

Grudin, J. and Pruitt, J. (2002). Personas, Participatory Design and Product Development: An Infrastructure for Engagement. *Proc. PDC 2002*. CPSR, pp. 144-161.

Guthrie, J. T. and Mosenthal, P. (1987). Literacy as Multidimensional: Locating Information and Reading Comprehension, *Educational Psychologist,* 22, pp. 279-297.

Hektor, A. (2003). Information activities on the Internet in everyday life. In: *New Review of Information Behaviour Research* 4/1 (2003), pp. 127 – 138.

Pruitt, J. & Grudin, J. (2003) Personas: Practice and Theory. In *Proceedings of the 2003 conference on Designing for user experiences*. San Francisco: ACM Press.

Macedo, A. A., Truong, K.N. & Camacho-Guerrero, J. A. (2003). Automatically Sharing Web Experiences through a Hyperdocument Recommender System. In: *Proceedings of the 14th conference on Hypertext and Hypermedia (Hypertext03)*, (Nottingham, UK, August, 26-30, 2003). Online: http://www.ht03.org/papers/pdfs/6.pdf [Available, Last checked: 6/16/2005]

Marchionini, G. M. (1995). *Information Seeking in Electronic Environments* .Cambridge: Cambridge University Press.

Marchionini, G. M. (2004). From Information Retrieval to Information Interaction, *Keynote at European Conference on Information Retrieval,* (2004, April). Online: http://ils.unc.edu/%7Emarch/ECIR.pdf [Available, last checked 5/9/2005].

Martzoukou, K. (2005). A review of Web information seeking research: considerations of method and foci of interest. *Information Research*, 10, Online: http://informationr.net/ir/10-2/paper215.html [Available, last checked 5/9/2005].

MIT (w.Y.): *The complete work of William Shakespeare.* Online: http://www-tech.mit.edu/Shakespeare/works.html [Available, last checked 5/9/2005].

Sinha (2003) Persona Development for Information-rich Domains. In: *CHI 2003 Conference on Human Factors in Computing Systems*. Fort Lauderdale, Florida.

Renear, A., Mylonas, E. & Durand, D. (1996). Refining Our Notion of What Text Really Is: The Problem of overlapping Hierarchies. In: N. Ide & S. Hockey (eds.): *Research in Humanities Computing*. Oxford University Press, Oxford, pp. 263-280.

Witt, A. (2005). Multiple Hierarchies: New Aspects of an Old Solution. In: Dipper, S., Götze, M., & Stede, M. (Eds.), *Interdisciplinary Studies on Information Structure (ISIS)*, Potsdam: Universitätsverlag Potsdam, pp. 55-86.

Witt, A., Goecke, D., Sasaki, F. & Lüngen, H. (2005). Unification of XML Documents with Concurrent Markup. In: *Literary & Linguistic Computing,* 20 (1), Oxford: Oxford University Press, pp. 103-116.