22

Open Corpus Adaptive Educational Hypermedia

Peter Brusilovsky[1] and Nicola Henze[2]

[1] School of Information Sciences, University of Pittsburgh
135 North Bellefield Ave., Pittsburgh, PA 15260, USA
peterb@pitt.edu
[2] IVS – Semantic Web Group, University of Hannover, & L3S Research Center
Appelstr. 4, D-30167 Hannover, Germany
henze@l3s.de

Abstract. Despite the fact that adaptive hypermedia techniques have proven their ability to provide user guidance and orientation in hyperspace, we do not currently see the widespread adoption of these techniques. A couple of reasons may explain this phenomenon. One of them is the current lack of re-usability and interoperability between adaptive techniques/systems, which – to some degree – originates in the so-called "open corpus problem" found in adaptive hypermedia. In this article, we analyze this problem in a popular arena: adaptive hypermedia systems with an emphasis on education. The origins and effects of the open corpus problem are discussed, and recent approaches are demonstrated that have – in one way or the other – developed as strategies for solving the open corpus problem. We summarize these findings and discuss how solution strategies can be successfully employed in the future, enabling adaptive hypermedia techniques within open, dynamic information spaces, such as the Semantic Web.

22.1 Introduction

The volume of educational resources available to students is changing rapidly. A variety of educational resources such as tutorials, electronic textbooks, and topic overviews are now available on the Web for almost every domain. Dedicated repositories of educational material, such as educational digital libraries (DL), and pools of reusable learning objects are being created. Finding high quality materials is much less of a problem with the use of modern Web search engines [6] and DL search services [43]. However, the resources that one finds have different presentation styles, target audiences, and coverage. Also, many resources are highly redundant. The abundance of resources has created another problem: How to help students find, organize, and use resources that match their individual goals, interests, and current knowledge? In brief, access is not the issue; *personalized access* is.

The need to provide personalized access to information is well recognized outside of education. Numerous research projects have proposed and investigated a wide range of techniques for personalized access within only the last few years. Earlier chapters of this book provided a good overview of personalization techniques for all major paradigms of information access: information retrieval in Chapters 6 [48] and 7 [47] of this book, browsing in Chapters 8 [8], and filtering/recommendation in Chapters 9 [56], 10 [52], 11 [59], 12 [16]. Successful application of personalization techniques has been achieved in such application areas as news access and e-commerce, covered in Chapters 18 [4] and 16 [28] of this book. Education, however, remains resistant to successful development while simultaneously being one of the few areas - accompanying medicine and public health - where the provision of personalized access is most important for users and society. The majority of adaptation techniques that focus on user interests and work successfully in other fields have a limited applicability in the educational context where users differ not just by their interests, but most essentially in their goals, skills, knowledge, and learning styles.

So far, the only techniques that demonstrate a good ability to provide personalized access to information in the educational context are adaptive navigation support techniques developed in the field of adaptive hypermedia (AH), presented in Chapter 8 of this book [8]. In a number of educational AH systems, adaptive navigation support techniques were able to help individual students locate, recognize, and comprehend relevant information, thus increasing learning outcomes and retention [7; 10; 13; 65].

Adaptive hypermedia techniques could provide a real difference for students who are trying to locate useful resources on the Web or in learning repositories and DL. Web resources rarely match the needs and the level of preparation of a specific class of students. Serious efforts are frequently required from students to understand which content is relevant, which is not, and how to find their way through it. Without individual guidance, students dealing with the increasing complexity of navigational possibilities may get lost in hyperspace in a number of senses. For example, they may fail to identify learning goals and recognize coherences, relations, and causal dependencies. Even in a learning repository where resources are carefully selected and classified by subject and category, the usefulness of resources depend on the individual learner's progress: some resources may require additional knowledge that the learner does not yet have (in accordance to his/her user model), while others may teach the subject without sufficient in-depth information and are thus too easy for this learner. At this juncture, methods from adaptive hypermedia can be used to support the learner in finding the *most appropriate* learning resource; for providing awareness about the learning process (e.g., by pointing out necessary pre-knowledge that this learner might otherwise miss); for providing guidance (e.g., by providing an individually tailored sequence of learning resources—teaching the topics s/he is interested in while incorporating all required prerequisite knowledge); for providing orientation (e.g., by pointing out the next learning steps to take, or the existence of different schools-of-thought); for considering individual learning styles; and so on.

Unfortunately, traditional adaptive hypermedia, with all its power, can't be directly applied in any of these important contexts. As it may become apparent from the study of Chapter 8 of this book [8], existing adaptive navigation support techniques are only able to work within a limited set of documents that have been manually structured and indexed with domain concepts and metadata at design time. Traditional adaptive hypermedia systems are predominantly *closed corpus* adaptive hypermedia, since the *document space* of these adaptive systems is a closed set of resources. Less than a handful of the adaptive hypermedia systems have attempted to deal with *open corpus* such as the Web's educational resources or dynamically expanding educational repositories. Closed corpus AH systems demonstrate what is possible to achieve with adaptive hypermedia technologies, but they are impractical for most real world applications because no teacher or content provider is able to invest time to structure and index thousands of documents collected from all over the Web as required by traditional adaptive hypermedia systems; worse, these systems would need constant maintenance, as new information becomes available daily.

The apparent contradiction between the potential power of adaptive hypermedia and its predominant close-corpus application content has caused a number of researchers to focus on what we call the *open corpus problem*:

> Is the applicability of the adaptive hypermedia techniques restricted by nature to closed corpus of educational resources or it is possible to develop *open corpus adaptive hypermedia* that will successfully work in such contexts as the Web and educational repositories?

The goal of this paper is to convince the reader that open corpus adaptive hypermedia is feasible and to discuss possible approaches to construct it. We start with a brief review of adaptation-specific information that is used by current adaptive hypermedia applications[1] (section 22.2). In Section 22.3 we stress several problems that have made it substantially difficult to use the current techniques of adaptive hypermedia with an open corpus of documents and review a range of known approaches and systems that attempt to overcome these problems, by attacking the open corpus problem from very different angles. Based on these considerations, we re-analyze the open corpus problem, especially with respect to the functional re-usability and interoperability of adaptive hypermedia and related systems (section 22.4). Then we discuss further, emerging solutions for realizing personalized access to information in open, distributed information spaces.

22.2 Adaptation-Specific Information in Adaptive Educational Hypermedia

To understand the essential difference between open and closed corpus adaptive hypermedia we want to start with more formalized definitions:

Definition 1 (Closed Corpus Adaptive Hypermedia System)

---

[1] We restrict ourselves to adaptive navigation support because we consider the first problem to solve in an open corpus setting is to adapt *access* to documents to the user's needs. *Content-level adaptation* (i.e., adapting the content of one particular document) is a possible future step which may later be addressed by, for example, considering information chunks or variations of documents instead of documents-as-entities, and by applying techniques from navigational-level adaptation onto these chunks / variations / etc.

A closed corpus adaptive hypermedia system is an adaptive hypermedia system which operates on a closed corpus of documents, where documents and relationships between the documents are known to the system at design time.

Definition 2 (Open Corpus Adaptive Hypermedia System)

An open corpus adaptive hypermedia system is an adaptive hypermedia system which operates on an open corpus of documents, e.g., a set of documents that is not known at design time and, moreover, can constantly change and expand.

What makes closed corpus hypermedia special, from the adaptation point of view, is exactly the fact that all documents and relations on the documents are known to the authors of an adaptive hypermedia system at design time. It allows the authors to augment the documents and relationships with additional information that can be used later by the adaptation algorithms to deliver the adaptation effectively to every user. We refer to this kind of information as *adaptation-specific* information. This information is typically hidden from the user; however, it is the real source of power of adaptive hypermedia. The goal of this section is to reveal the kind of information that is used by adaptive educational hypermedia in order to perform the adaptation. Understanding the nature adaptation-specific information can help us to identify the information necessary for the open corpus context, and how to produce or compensate for this information.

It turns out that we can distinguish two classes of this adaptation-specific information: The first class comprises information that adds some semantics to the hypermedia, i.e., assigns specific types to the hypermedia documents and relationship and introduce additional, semantic relationships between the documents. The second class provides additional knowledge "behind" the hyperspace documents by connecting documents to external models that are separate from the hyperspace itself. The variety of these models is high: we can find conceptual models, pedagogical models, goal models, stereotype hierarchies, and more. Many approaches from artificial intelligence have been used to verify, maintain and interpret these models in order to perform the adaptation task for these adaptive educational hypermedia systems.

22.2.1 Adaptation-Specific Information: Enriching Hypertexts with the Annotation of Documents and Relationships

The first kind of annotation-specific information acts within the hyperspace itself, attempting to introduce some additional knowledge about the documents (nodes), links, and additional relationships between the documents. Additional knowledge about the documents and links is typically provided in the form of *types* assigned to documents and links. For example, in the KBS-Hyperbook [34], which is presented in more detail in section 22.3.2, documents can be marked as "problem statement," "example," "theory," etc. Links are usually typed to reflect the semantics of the structural relationship between connected documents. Some systems use an elaborate set of typed links [2]; however, other systems such as MetaLinks [51] achieve a good functionality while using just two types of links. From the modern point of view, link and document types can be seen as metadata that is added to documents. However, in adaptive educational hypermedia systems, they are normally referred to not as metadata, but as knowledge about documents and links. This knowledge about documents connected to the current document, and about the connection types, allows the adaptive hypermedia decision mechanisms to guide the user to the most appropriate documents, using such techniques as link ordering, annotation, and hiding presented in Chapter 8 of this book [8].

In adaptive educational hypermedia systems, the typing of existing links is often not sufficient, because these systems rely on knowledge dependencies or pedagogical relationships between documents that may not be directly connected by a link. To compensate, these systems introduce additional relationships between documents that are often invisible, i.e., not accessible for navigation. Most typical among these relationships is a prerequisite relationship that notes that one document should be known before another. It is used in many systems such as ELM-ART [64] or AHA! [22]. Document-to-document relationships are very powerful in adaptation. On the other hand, the drawback is that alterations to the set of documents in hyperspace normally requires a huge effort. Consider, for example, the introduction of a new document: At which points shall it be presented to learners? Which documents are prerequisites, and to which documents is this document a prerequisite? Checking all documents of the documents space one-by-one may be required to establish proper relationships.

22.2.2 Adaptation-Specific Information: Connecting Hypertexts with External Models

An alternative way to add adaptation specific information is to rely on external models that exist beyond the hyperspace, such as knowledge models, pedagogical models, usage models, etc. These models typically encapsulate some kind of knowledge. For example, conceptual domain models encapsulate knowledge about the domain while stereotype hierarchies represent knowledge about users. In this case, the necessary knowledge is added to hypertext documents by connecting them to elements of these external models. Most popular in the field of adaptive educational hypermedia are the domain concept models, presented in Chapter 1 of this book. The use of domain concept models, along with user overlay models, allows these systems to provide sophisticated adaptations to the user's level of knowledge. For example, in the InterBook system [11], the authors connect (or *index*) documents with domain concepts using two kinds of document-concept relationships – the outcome and prerequisites. These links allow the authors of the system to express what domain knowledge is presented in the page or what knowledge should have been mastered before the page is accessed. Other models, such as didactical models, provide information on a certain didactical approach, and can be seen as a new layer to both the document-to-document annotation (internal references) and the document-to-concept annotation approach (external references). Generally, storing adaptation-specific information in external models supports the application of artificial intelligence techniques for reasoning about this information.

Indexing documents in terms of external models provides for a higher level of adaptation than simply typing and connecting documents, since these models typically encapsulate additional knowledge that can be used by adaptation algorithms. However, it also adds an additional challenge to the system development since the building of sound external models is a considerable knowledge engineering effort that typically requires expert knowledge in a specific field. The initial investment into developing external models pays off, to some extent, since this allows the indexing of documents to become easier: authors can write their materials, and index it with concept models without considering the whole set of *currently* available documents. In particular, multi-author approaches are supported, where material can be designed and annotated independently. Another advantage of document-to-concept relations are achieved with respect to *maintenance*: changes in the document space affect only the altered / added documents, no further annotations of documents need to be altered.

22.3 Several Ways to Open Corpus Adaptive Educational Hypermedia

How can we achieve progress in developing open corpus adaptive educational hypermedia systems that are compatible in personalization power with existing closed corpus systems? Arguably, this goal can be achieved if we find the way to enhance open corpus resources with additional knowledge that is comparable with the knowledge behind traditional adaptive hypermedia, which was analyzed in the previous section. If comparable knowledge could be obtained in the open corpus context, existing adaptive hypermedia techniques or their modifications could be used to deliver adaptive navigation support for open corpus documents. This section attempts to analyze several known ways of developing open corpus adaptive hypermedia, i.e., several ways to collect the missing knowledge. Following the structure of the previous chapter, we separate the discussion into two subsections - one dealing with intra-hyperspace problems such as document interlinking and link typing and the other dealing with the problem of external models and the indexing of hyperspace documents.

22.3.1 How to create a linked hyperspace from open corpus resources

Adaptive hypermedia technologies support hypertext-browsing activities of the user, i.e., they assist the user in moving from document to document, following inter-document links. Thus, the first problem to resolve when building an open corpus adaptive hypermedia system is how to build a hyperspace from an open collection of generally independent documents. This problem could be solved in two different ways: relying on human power to create hyperspace and creating a hyperspace automatically.

As discussed below, the manual interlinking of a constantly expanding set of open corpus documents is possible in only a few contexts. A more general solution is to apply some techniques that can automatically create a linked hyperspace from a collection of independent documents. The problem of automatic hyperspace creation has been explored by researchers in the area of hypertext and information retrieval for at least 20 years. This research started originally under the term *intelligent hypertext* and focused mostly on automatic linking of documents as a help for hypertext authors who may not be able to identify all useful links. Later, the ideas of automatic linking were explored by the *open hypermedia* movement. Open hypermedia, as a research direction, specifically focused on conceptual and architectural problems of creating a hyperspace from originally independent open corpus documents [50]. A large body of literature has been produced in both areas and a range of techniques has been suggested. These techniques can be generally classified into two groups that we call keyword-based techniques and metadata-based techniques.

Manually Constructed Hyperspaces. Relying on human power to create hyperspace is a possible solution for an open corpus system, so long as the developers of this system are not involved in the hyperspace construction. In fact, the simplest way to explore open corpus techniques for educational AH is to take an existing educational hypertext application and to add a layer of adaptive navigation support to it. This approach allows the developers of open corpus adaptive hypermedia system to simply avoid the problems of hyperspace construction and focus on navigation support techniques. A number of early explorers of open corpus educational AH have used this approach with various kinds of pre-existing hypermedia applications, such as an educational encyclopedia [36], a hypertext tutorial [32], and an educational Web site [58]. Two of these systems are presented in more detail in section 22.3.2 of this chapter.

While early projects operated in the context of pre-authored hyperspace, similar approaches could be applied to the constantly expanding yet human-linked document collections. The challenge here is to find an environment where the human-supported hypertext construction is supported naturally, i.e., where each new resource is being immediately linked to the whole collection by a human author or manager. While in most of the cases this is not feasible, there are at least two meaningful contexts that deserve further exploration. One context is organization-supported hyperspaces such as Web sites or educational portals where the integration of new resources into the previously linked hyperspace is ensured by the organization that owns or maintains the collection. Unfortunately, this context is becoming more and more rare: due to high cost of manual linking, many portals and resource collections adopt a *pool approach* where each new resource is simply added to the pool. Another context is *community-driven* hypertext creation, where linking new documents to the existing hyperspace is done by a whole community of users. Two popular examples of community created hypertexts are Wikis and blogs where the nature of these community-based systems encourages linking newly authored documents. Both kinds of expanding hypertext systems, organization-supported and community-driven, provide a really challenging but creative application area for open corpus adaptive hypermedia and we expect more work in this direction in the coming years. A pioneer example of open corpus adaptive hypermedia for Wiki is the CoWeb system [24], which uses the ideas of *social navigation* to provide annotation-based adaptive navigation support. CoWeb is briefly reviewed in sectioin 22.3.2 below.

Keyword-based Techniques for Automatic Hyperspace Creation. This group of techniques is based on the automatic keyword-level analysis of documents. The work on keyword-based linking started at the end of 1980 with exploring similarity-based navigation. The idea of similarity-based navigation is to create links between documents that are similar on the keyword level. The techniques for calculating keyword-level similarity are well explored and covered in more detail in Chapter 5 of this book [49]. Since the pioneer work of Mayes and Kibby [40; 46], keyword level similarity-based navigation has been applied in a number of systems. This automatic linking technology is simple and straightforward and can be used in almost any context. To interlink an existing collection of documents, a similarity metrics is calculated between each pair of documents. To link a new document to a collection, the similarity is computed between the new document and all documents in the collection. After that, documents with similarity higher than a certain threshold are connected by a bi-directional hyperlink.

The negative side of this technology is that the quality of simple keyword-level similarity techniques is not perfect, so it can often link pages that are not really semantically related. In addition, a hyperspace created with classic similarity-based navigation techniques suffers from two problems – the lack of typed links and the lack of clear structure (the resulting hyperspace is rather chaotic). As we observed in section 22.2, typed links enable more advanced navigation support technologies. A clear hypertext structure helps users to find their way and position in hyperspace.

More recent research attempts have focused on overcoming these problems using more advanced keyword-based techniques. The first challenge to be addressed was link typing. By the end of 1990, a number of keyword-level techniques were suggested for generating typed links [2; 21] as well as typing existing hypertext links [1; 54]. Several researchers focused on improving the precision of keyword-level linking by replacing standard document indexing with "semantic-oriented" techniques such as latent semantic indexing [44] or lexical chaining [29].

To structure a collection of unrelated documents, several researchers applied Self-Organized Maps (SOM). The SOM technology is able to cluster documents into cells on a rectangular grid in such a way that documents allocated to the same cell are quite similar to each other and documents in the neighboring cell are also similar, although to a lesser extent [41]. This unique property of SOM allows the introduction of some reasonable level-structuring even in a large collection of Web resources [20; 42; 55]: each cell or group of cells serves as a category (section of hyperspace) with spatial proximity expressing similarities between the categories. Thus the application of SOM turns a collection of documents into a structured spatial hypertext (spatial hypertext implies implicit links between spatially co-located documents [57]). Using map-based navigation, introduced in [14], this spatial hypertext can be converted into a regular hyperspace that allows navigation from a document to the hosting map cell and then to similar documents. This technology has been applied in the Knowledge Sea II system, presented below.

Metadata-based Techniques for Automatic Hyperspace Creation. Another branch of research that may resolve problems of simple keyword-level hypertext linking is the application of metadata-based techniques. Generally, metadata-based approaches allow the production of better quality results in the linking and structuring of hyperspace. The early focus on keyword-level techniques was justified by the lack of metadata. However, over the last several years a number of repositories have assembled a large volume of documents indexed with metadata. In addition, some progress has been achieved by extracting metadata from Web resources. As a result, metadata-based approaches now overshadow keyword-based approaches. With the exception of link typing (which is not a problem in the presence of metadata), the work on metadata-based hypertext construction has been focused on the same goals – automatic linking and automatic structuring.

The pioneer work on metadata-based linking was done in early 1990 by the team of Douglas Tudhope [60]. They explored similarity-based navigation in a richly metadata-indexed photo-archive. The core idea of similarity-based navigation is the same as for keyword-based linking: a similarity measure is computed between documents and those with similarity above a certain threshold are connected by a link. The metadata similarity is calculated as a weighted measure of similarity along each metadata facet. This process is presented in more detail in Chapter 11 of this book [59]. Since metadata expresses semantic similarity (in contrast to surface similarity expressed by keywords) this approach obtains high quality links. More recently, the focus of research on metadata-based linking moved from simple quantitative metadata (such as time, size, or difficulty) to ontology-based linking. Ontology-based linking is possible when documents are indexed (manually or automatically) with terms of ontology or a thesaurus. In this case, the process of finding similar documents is a more challenging process, since it has to take into account the position and connections of ontological tags in the ontology. A well-known approach to ontology-based linking in open corpus hypermedia is presented in [19]. Currently these ideas are being explored in the context of the Semantic Web [25].

The presence of metadata also allows some meaningful structuring of hyperspace. The complexity of possible structuring is determined by the complexity of metadata indexing.

- The simplest case of metadata indexing (single-facet non-ontological metadata) allows the *grouping* of documents that share the same metadata value and the organization of *concept-based navigation* between independent documents [15]. Concept-based navigation is based on a set of additional navigation "concept" pages – one for each value of metadata (for example, the author of a publication). Each of the concept pages provides links to all documents indexed with this value. Each document is also connected to concept pages corresponding to all concepts from its index. Concept-based navigation allows a user to navigate from a document to any of the related concept and then to any other document indexed with the selected concept.
- The presence of ontological metadata allows organizing documents into a hierarchy (which is known as the best browsing framework) along the structure of the ontology used for indexing. The user can navigate the collection of documents along the ontology tree where a visit to each node (taxon) of the tree (or, at least, of each terminal node) provides access to all documents indexed with this taxon. The

user can also use an extended version concept-based navigation moving from a document to a concept related to this document, that to a concept connected to the first concept in the ontology, and then to a document connected to the second concept. This powerful navigation approach is currently used in many resource repositories and has already been considered for open corpus adaptive hypermedia [45].

- Finally, the presence of faceted ontological indexing (multi-faced indexing with ontological metadata) allows to generation an exceptionally rich lattice-based navigation structure. This case (now typical for many digital libraries) further extends ontology-based navigation with an opportunity to navigate along several taxonomies (switching them on the fly). A good example of using navigation opportunities provided by multi-faced ontological indexing is the Flamenco browser [67].

22.3.2 External Models and the Indexing of Open Corpus Resources

As explained in section 22.2, one of the keys to providing adaptive navigation support is the presence of knowledge (adaptation-specific information) behind documents. While document and link typing provides us with some knowledge, in classic adaptive hypermedia this knowledge is most frequently provided in a different way: by connecting documents to *external models* – such as domain models, pedagogical models, or stereotype hierarchies. This process is known as *indexing*. More information about it is provided in Chapter 8 of this book [8]. Both, the creation of external models and indexing are traditionally done by the authors of adaptive hypermedia. To apply comparable methods to the open corpus of documents, one needs to resolve two related problems: where to find external models and how to index open corpus documents in terms of these models.

Existing open corpus adaptive hypermedia systems have explored several ways to solve these problems. Quite similar to the case of hypertext construction and typing, we can distinguish between manual and automatic technologies, with the automatic technologies being classified as keyword-based, metadata-based, or community-based.

The Manual Indexing of Open Corpus Resources. The manual indexing approach assumes that all adaptation-specific information is added to documents by humans, although not by the system developers and possibly after the core system has been created. This approach was explored in a few classic adaptive hypermedia systems that attempted to cover open corpus documents: KBS-Hyperbook [34] and SIGUE [18]. These systems used the classic kind of external models – domain models in the form of a network of concepts (see Chapter 1 of this book [12];) however, their corpus of documents was not closed. Any Web document could be integrated into the systems as soon as it is indexed with domain concepts. This approach is limited in its applicability yet it can be used in expanding document repositories where the manual indexing of incoming documents is feasible. The positive side of this approach is the ability to use high-level models: it results in a good quality of indexing. When coupled with concept-level overlay user models, presented in Chapter 1 of this book

[12], manual concept indexing could support most advanced navigation support techniques. The negative side is the high price of model development and indexing.

A representative example of the manual approach is the KBS Hyperbook system. The first prototype of the KBS Hyperbook was developed in 1998 [32]. The fundamental concept behind the adaptation component in the KBS Hyperbook system is the separation of the adaptation module from the hypermedia system itself. This is realized by a rigorous separation of the reasoning engines and the resulting adaptation functionality from the module for organizing and maintaining the hypertext structure.

The KBS Hyperbook uses an external domain model, which serves various purposes. First of all, the domain model describes the application domain by defining all concepts relevant to the domain, as well as the relationships between these concepts. The domain model is created manually, and is the only source of knowledge about the domain that the system uses. Secondly, the domain model's concepts are used to link the hyperspace to this external domain model. Thirdly, the domain model provides the main source for the creation of a Bayesian Network [53], whose main responsibility is to estimate the actual knowledge state of a user U at any given time.

The indexing of hypermedia documents covers two dimensions: the first dimension describes the content of the document. This is done by indexing each document against a set of concept names from the external domain model. In addition, this indexing step also provides the necessary information for linking the hyperspace to the external model. The second dimension contains attributes that state the type of document, referring to a so-called conceptual model. The conceptual model defines possible types of documents found in this domain, e.g., such categories as "problem statement," "example," "theory," etc.

Thus, if a document is added or modified, the author has to assign a set of concepts describing the type and content of the document (or modify these attributes accordingly). In this way, the system meets the requirement that the metadata annotations for the document are independent from the application domain.

KBS Hyperbook is therefore an early showcase of open corpus adaptive hypermedia, relying on external models (domain model / Bayesian Network, and conceptual model) and the metadata arising from them, which is then added to the hypermedia documents. Several adaptive e-learning systems have been realized with the KBS Hyperbook technology, the most prominent among them being the Hyperbook on Java Programming, the *Java Hyperbook*. The Java Hyperbook is an adaptive system which uses course materials from an undergraduate course on Java Programming held at the University of Hannover and guides the student through the course by showing the next reasonable learning steps, selecting projects, generating and proposing reading sequences, annotating the educational state of information, and then by selecting information that will be useful to the user, based on their actual goals and knowledge [33]. To prove the openness of the Java Hyperbook, the authors added the content of the Sun Java Tutorial [17], a freely available online tutorial, to the Java Hyperbook. The Java Hyperbook was capable of adapting to both corpora [34]. A screenshot of the Java Hyperbook is displayed in Figure 22.1.

However, the coding of learning dependencies into the external domain model has been shown to be a drawback to the simultaneous integration of different corpora: Each collection of learning materials may follow their own learning / teaching strategy and therefore may define different learning dependencies, resulting in a different structure to the domain model. Thus, the coding of knowledge in an external domain model has the clear advantage of making the indexing and integrating of new documents is very straightforward and cost efficient. In addition, when the chosen domain model can be applied successfully to these new documents, the adaptation will work immediately. This approach is functioning very well for the Java course and Sun Tutorial, since both share the same view of the domain. In cases where this constraint is not given, the adaptation may not work correctly. Thus, the KBS Hyperbook approach is applicable to an open corpus of documents where each of the corpora shares a common domain model view.

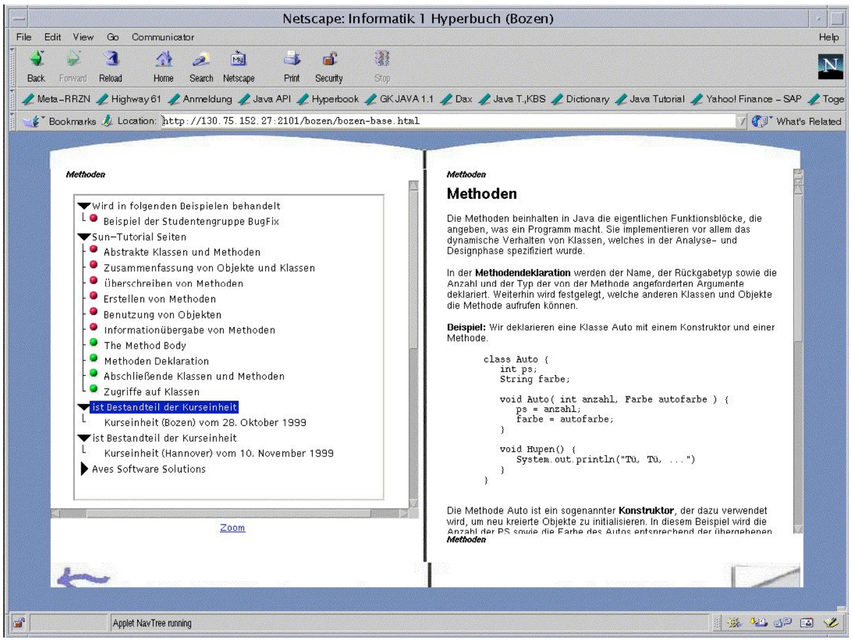

Fig. 22.1. Example of the KBS Hyperbook for Java Programming, displaying a learning unit on methods in Java (on the right hand side). The left side is composed of links to relevant learning material. Traffic light annotations of the links recommend to each learner certain navigational possibilities over others. The Sun Java pages are enriched with these recommendations, matching the previously annotated materials from the Java course corpus. Along the top, there is an array of references links, e.g., to examples, references, the Sun Java tutorial, and courses where this learning unit is used.

To overcome this drawback, the KBS Hyperbook team continued with a focus on the following issue [35]: A generic knowledge modeling approach for adaptive, open corpus hypermedia systems based on ontology modeling. For each corpus of documents integrated into the open adaptive hypermedia system, a sub-graph of the ontology was calculated, with the goal of estimating the user's knowledge with respect to this corpus. This enables the KBS Hyperbook to maintain different domain models (corresponding to these sub-ontologies), which are related to each other via the common overall ontology.

Automatic Keyword-Based Indexing. In contrast to the manual approach, automatic keyword-based methods offer a low-price solution. These methods use the information retrieval approach to document modeling that is presented in Chapter 5 of this book [49]. The role of the external model is played here by a set of meaningful keywords. This is a relatively low-level model, however it can be automatically produced by document analysis and supports simple, automatic indexing of documents. Coupled with keyword-level user profiles, presented in Chapter 2 of this book [27], this approach is used in the majority of content-based recommendation systems (see Chapter 10 of this book [52] for more details). This approach has been successfully used in a number of contexts to recommend open corpus resources that are relevant to user *interests*, but its ability to adapt to other user aspects – such as knowledge or goals is very limited. Another negative side of this approach, from the educational viewpoint, is its lower precision. This reduces its applicability to educational context where adaptation to knowledge and learning goals is typically more important than adaptation to interest and where reliability of guidance is critical—because students typically aren't capable of judging how relevant an educational resource is.

An attempt to apply the keyword-level approach in adaptive educational hypermedia was done in the ML Tutor system [58]. The ML Tutor is a hypertext system that provides suggestions to the user on the basis of their recent browsing history, indicating pages that are relevant to the user's current area of interest. The ML Tutor was specifically designed to support user navigation in Web-based hypermedia, however, its internal mechanism is not essentially different from *Syskill and Webert* and other "generic" content-based recommender systems presented in Chapter 10 of this book [52]. Like many of these systems, the ML Tutor also applies machine-learning techniques to "learn" the user's profile of interests by observing browsing behavior and then recommends the most relevant pages in ML Tutor's "known" part of the Web.

The hyperspace used by ML Tutor is constructed manually, but not by the authors of the system. Instead, the authors integrated four existing independent Web sites, connecting them with additional "bookmark" links to form a joint hyperspace with 133 nodes. The role of the domain model in ML Tutor is played by a list of domain-specific keywords that were constructed manually. The indexing process is fully automatic. The result of this indexing is a binary vector for every page, where each keyword that is present in the page is indicated by 1 and each absent word by 0. The page vectors are stored in the internal database along with page IDs and URLs.

The system is implemented as an applet communicating with the server-side Machine Learning Component (MLC) of the system. At runtime, the ML Tutor applet passes the URL addresses of the last ten hypertext pages visited by a user to the MLC. Knowing their page vectors, MLC produces a list of recommended hypertext pages that focus on the same "topic" but are not yet visited and sends this information to the ML Tutor applet. The list of recommended links is displayed to the user in typical recommender-system style, with non-contextual link generation, as shown in figure 22.2.

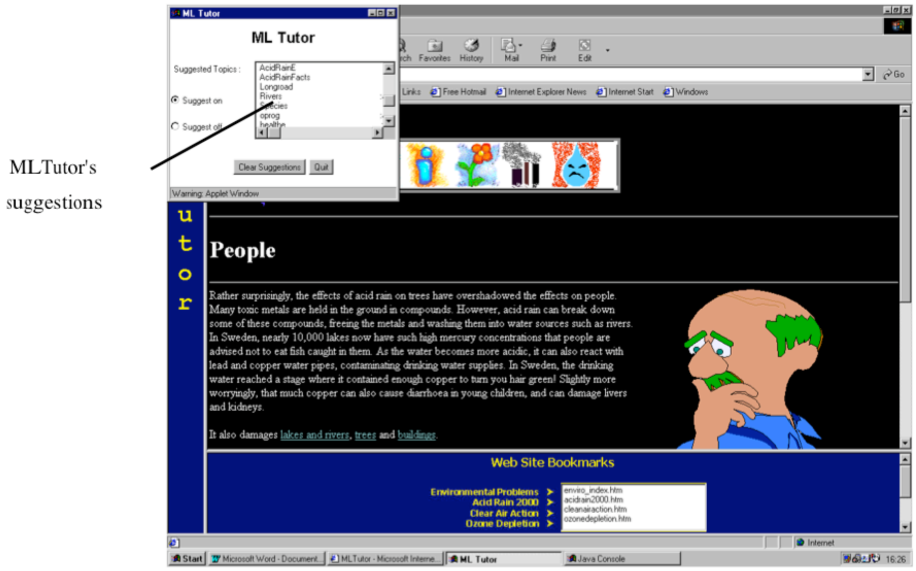

Fig. 22.2. Adaptive navigation support using non-contextual link generation in the ML Tutor. The list of suggested pages is shown in a separate window in the upper left corner. The figure is reused from [58] with the publisher's permission.

Metadata-Based Approaches. The goal of these approaches is to improve the quality and range of supported adaptation techniques by using higher quality semantic knowledge about a document. When metadata is added to a document, it may provide important information about the document's content, intended use, primary reader group, difficulty, etc. In the area of e-learning, these metadata-based approaches benefit from the existence of standards for describing learning resources (or so-called Learning Objects). An example of an adaptive educational hypermedia system that makes use of metadata is the Personal Reader Framework (PRF) [30], which provides an environment for designing, maintaining, and running *personalization services* in the Semantic Web. The goal of the framework is to establish adaptation functionality as a Semantic Web service, which can be encapsulated and re-used.

In the run-time component of the framework, Personal Reader instances are generated by plugging together one or several of these personalization services. Each developed Reader consists of a browser for learning resources (*the reader part)* and a side-bar or remote, which displays the results of the personalization services, e.g., individual recommendations for learning resources, contextual information, pointers to further learning resources, quizzes, examples, etc. (*the personal part*). A screenshot of a Personal Reader for learning the Java programming languages is depicted in Figure 22.3.

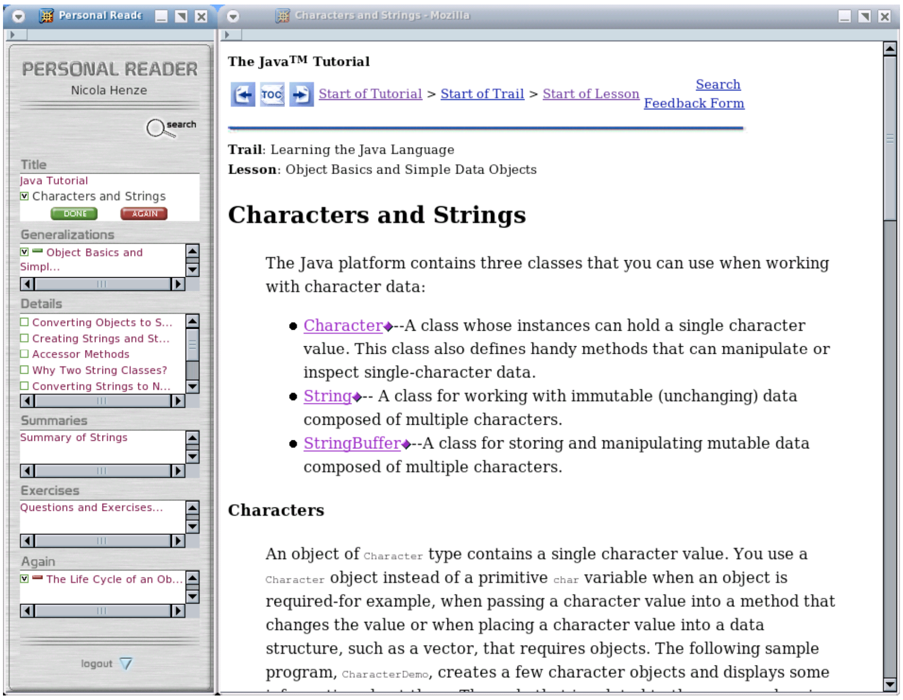

Fig. 22.3. Screenshot of the Personal Reader for Java programming The Personal Reader consists of a browser for learning resources *(the reader part)* and a side-bar or remote, which displays the results of the Personalization Services, e. g., individual recommendations for learning resources, contextual information, pointers to further learning resources, quizzes, examples, etc. *(the personal part).*

The PRF makes use of recent Semantic Web technologies for realizing the service-based environment necessary for implementing and accessing personalization services. The core component of the PRF is the so-called *connector service* whose task is to pass requests and processing results between the user interface component and available personalization services, and to supply user profile information, and available metadata descriptions on learning objects, courses, etc. In this way, the connector service is the mediator between all services in the PRF.

Two different kinds of services - apart from the connector service – are used in the PRF: personalization services and *visualization services*. Each personalization service offers some adaptive functionality, e.g., recommends learning objects, points to more detailed information, quizzes, exercises, etc. Personalization services are available to the PRF via a service registry using the WSDL (Web Service Description Language, [63]). Thus, service detection and invocation take place via the connector service, which asks the Web service registry for available personalization services, then selects appropriate services from this list. The task of the *visualization services* is to provide a user interface for the Personal Reader: interpret the results of the personalization services to the user, and create the actual interface, composed of reader and personalization sections.

The PRF refers—as far as possible—to standard metadata annotations: The currently implemented sample readers (for the domains "Java Programming" and "Semantic Web") make use of metadata descriptions for documents in accordance with LOM [38], while user profile information relies on the IEEE PAPI specification for describing learners [37]. Further, domain ontologies are applied: e.g., domain ontologies for Java programming or the Semantic Web. By using ontologies for describing run-time user observations and for adaptation, these models can be shared with other applications. The PRF can also implement concurrent personalization services which fulfill the same goal (e.g., provide personal recommendations for some learning object), but which consider different aspects in the metadata. For example, one personalization service can calculate recommendations based on the structure of the learning materials in some course and the user's navigation history, while another checks for keywords which describe the learning objectives of that learning object and calculates recommendations based on relationships to the corresponding domain ontology. Examples of such personalization services are described in [30].

The Community-Based Approaches. *Community-based* approaches to open corpus adaptive navigation support are based on the idea of *social navigation*. Social navigation tries to solve the navigation problem by taking advantage of the natural human tendency to follow the footsteps of other people with similar interests. Similar to collaborative filtering systems (see Chapter 9 of this book [56]), these approaches ignore the *content* of the documents, relying instead on information about the *usage* of these documents by a community of users. In a community-based approach, a document is "indexed" with all users who paid attention to this document explicitly or implicitly (i.e., rated, read carefully, bookmarked, or printed it). Thus, the community (or communities) of users of the system serves as an external model.

The CoWeb system [23; 24] mentioned in the section 22.3.1 above provides a good example of a simple social navigation system that works in manually authored hyperspace or a Wiki system CoWeb. To increase awareness of what is going on in the CoWeb and to guide the users to most recently updated or visited pages all links inside the CoWeb were annotated with activity markers (Figure 22.4). An *access marker* showed access information using a metaphor of *footprints*. Small footprint symbols in three different colors (gray, orange, red) were placed right next to links to indicate the amount of traffic the page behind that link received in the past 24 hours. A *novelty marker* also in three different levels indicated how long ago that page was last modified.

A more sophisticated example of social navigation support is the Knowledge Sea II system [9], which attempts to automate both hypertext construction and indexing. Knowledge Sea II relies on SOM technology for the hypertext construction, which was introduced by its predecessor system Knowledge Sea [14]. Knowledge Sea applied SOM to build an 8 by 8 knowledge map from several thousands of Web pages belonging to several independent online resources for learning C programming language. As was mentioned earlier, SOM technology allowed the placement of similar pages into the same or adjacent cells on the map. Using *map-based navigation* [14], the users of the system were able to navigate from a page to the cell it belongs to, to connected cells (if necessary), and finally to pages that were similar to the page where they began their map-based navigation.

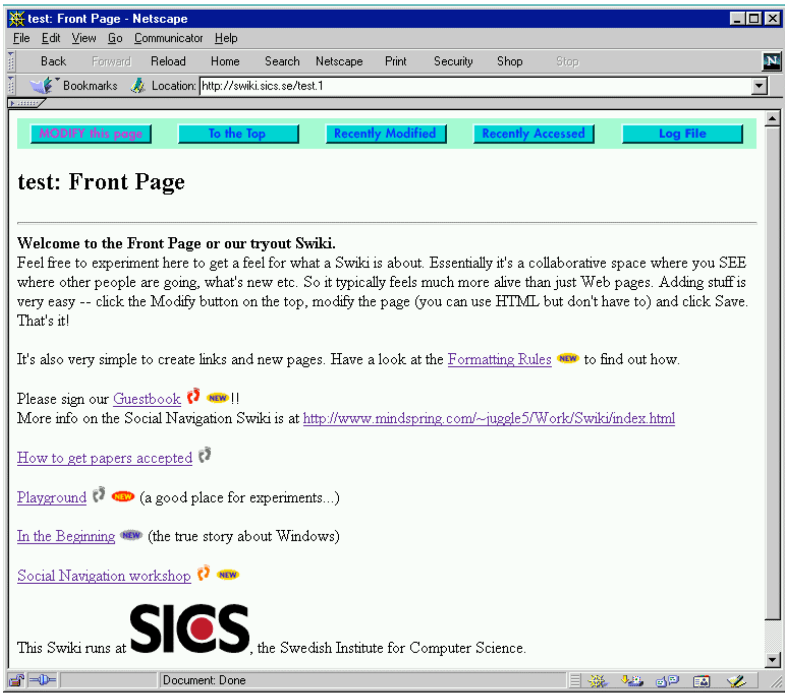

Fig. 22.4. Social navigation support in CoWeb. Two kinds of activity markers indicate when the page behind the link was last modified and also whether it was recently accessed. Used from [23] with the permission from the author.

Knowledge Sea II expanded the original Knowledge Sea with two kinds of adaptive navigation support, based on social navigation concepts: traffic- and annotation-based navigation support. Both kinds of navigation support are provided by generating visual cues that change the appearance of links on the pages and map cells presented to the user (Figure 22.5). The system generates appropriate cues individually for each user by analyzing past individual activities of the user and other users belonging to the same group.

Traffic-based navigation support attempts to express how much attention the user herself and other users from the same group paid to each of 25,000 pages that the system monitors. The level of attention for a page is computed taking into account both number of visits and time spent on the page and is displayed to the user through an icon that shows a human figure on a blue background. The color saturation of the figure expresses the level of the user's own attention while the background color expresses the average level of group attention. The higher the level of attention is, the darker the color appears to the user. The contrast between colors allows the user to compare her navigation history with the navigation of the entire group. For example, a light figure on a dark background indicates a page that is popular among group members but remains under-explored by the user. The color of the map cell and the human figure shown in the cell is computed by integrating attention parameters of all pages belonging to that cell.

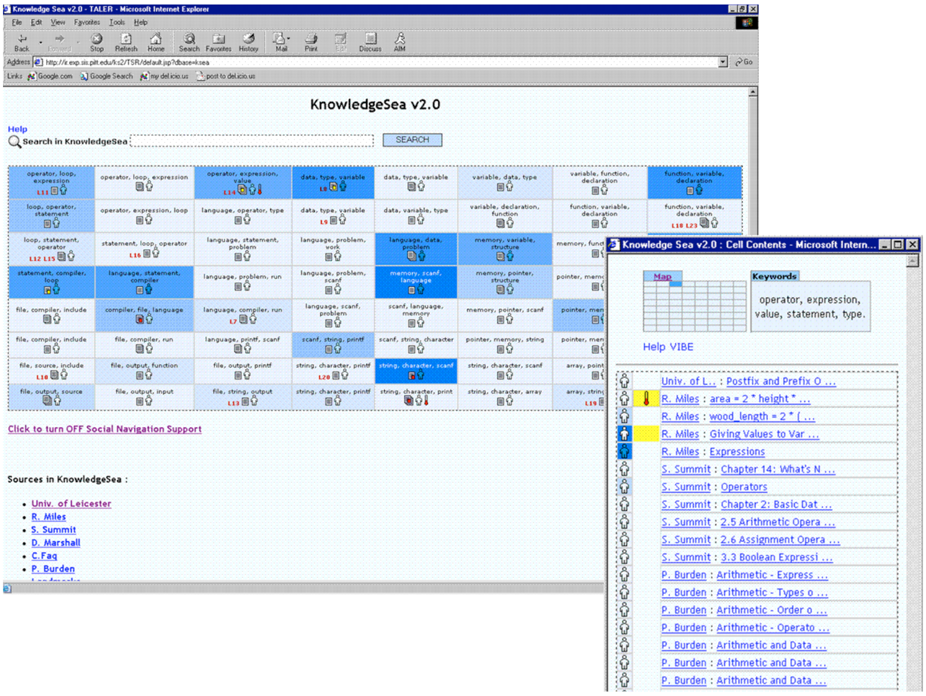

Fig. 22.5. Social navigation support in the Knowledge Sea II system. The knowledge map is shown on the left and an opened cell on the right. A darker blue background indicates documents and map cells that have received more attention from users within the same group. Human icons with darker colors indicate documents and cells that have received more attention from the user herself. Similarly, a yellow background indicates density of annotation and a thermometer icon measures how positive these annotations were.

Annotation-based navigation support uses a similar approach to represent the number of page annotations made by the user herself and other users from the same group. Each page in the system can be annotated by the user. The user can also indicate that this note is praise (i.e., the page is good in some aspect). While users make annotations mainly for themselves, Knowledge Sea II allows all users of the same group to benefit from collective annotation behavior. The yellow annotation icon shown next to the blue traffic icon shows the density and the "praise temperature" of annotation for each page. The more annotations a page has, the darker the yellow background color appears to the user. The temperature shown on a thermometer icon indicates the percentage of praise annotations.

Knowledge Sea II provides a good case for stressing the positive and negative sides of the use of community-based approaches with implicit feedback for adaptive navigation support. On the positive side, Knowledge Sea II requires the least effort to add a new document to the system: neither manual nor automatic page pre-processing is required for the navigation support part to work (however, note that automatic processing is required to add a page to the map since the system uses a keywords-based hyperspace creation approach). As a result, Knowledge Sea II can instantly add any new document to the system as soon as the first of its users encounters it during navigation. This gives it a ranking as the most "open" of all the approaches to open corpus adaptive navigation support. On the negative side, the navigation support provided by community-based technologies is relatively weak and is sensitive to the system's ability to identify a group of "similar" users.

22.3.3 Discussion

The previous sections demonstrate the existence of a whole range of approaches that might be able to overcome two aspects of the open corpus problem in adaptive educational hypermedia. It's interesting to note that both the existing hypertext construction and page indexing approaches can be grouped into four similar categories – manual, keyword-based, metadata-based, and community-based (Table 22.1). While community-based hypertext linking approaches have not been analyzed, they do exist [5; 66]; the authors have simply failed to find an example of these approaches, used in an appropriate context.

Table 22.1. A Summary of hyperspace construction and document indexing approaches for the open corpus, with examples of actual systems

Approaches	Hyperspace construction	Document indexing
Manual	KBS-Hyperbook[32]; MT Tutor [58]; CoWeb [24]	KBS-Hyperbook [32]; SIGUE [18]
Keyword-based	Knowledge Sea [14]	MT Tutor [58]
Metadata-based	COHSE [19]; Flamenco [67]	PRF [30]
Community-based	Bollen & Heylighen [5]	CoWeb [24]; Knowledge Sea II [9]

It is important to stress again that these approaches do not contradict but rather complement other since they have different strong and weak sides. So far, a number of existing systems have combined different approaches to hypertext construction and navigation support. For example, Knowledge Sea II uses the keyword-based approach to create the hyperspace and a community-based approach to provide navigation support. However, it is certainly wise to combine different approaches to achieve the same goal – as has already been done by the hybrid recommender systems presented in Chapter 12 of this book [16]. Moreover, an interesting challenge is to integrate approaches so that they will support each other. For example, the techniques used for analyzing social navigation patterns or for identifying Web communities may be used to help detect hidden relations between documents, where these relations might express similarity in content, as well as finding contradicting relations between documents and others. So, these techniques can be used to gather metadata—based on usage, structure, or content—for the hypermedia components, the hypertext documents, and hypertext relations. This metadata can then be used within a personalization service for recommending and visualizing information, as in Personal Reader. Vice versa, the metadata used in the Personal Reader can be used to further strengthen the pattern-detection algorithms of the social navigation process.

22.4 The Road Ahead

In this section, we analyze the effects of the open corpus problem on reusability and interoperability issues. As a conclusion to this, we will discuss the open corpus problem in relation to the Semantic Web, and give possible solutions for overcoming the open corpus problem and its implications. Although our starting point has been the open corpus problem in the field of adaptive educational hypermedia, many of the below considerations are valid for adaptive hypermedia in general.

22.4.1 Re-usability and the Open Corpus Problem

Traditional adaptive hypermedia systems operate on some fixed document space, where documents and relations between them are known at the design time and adaptation strategies are developed with respect to this specific set of documents. Especially document-to-document relations (see section 22.2.1) can only be validly assigned if the complete document space is known. Adaptation algorithms deliver faulty results if the document space is altered (e.g., if documents are modified, deleted, or new documents are introduced) as the document-to-document relations used in the algorithms become invalid. Only sophisticated re-engineering of the metadata (again on the complete document space) can recover the situation. One implication of the closed corpus in traditional adaptive hypermedia is that adaptive applications consequently fail in exchanging content with other (adaptive or non-adaptive) applications. Thus, the *re-use of content*—a very important aspect, especially when it comes to the Web—is not supported. To achieve re-usability, substantial re-engineering of particular systems is required, which cannot be realized in an on-demand basis.

In the context of e-learning, recent developments have yielded not only metadata standards for e-learning but also large collections / repositories of learning material, where both learners and teachers can store and retrieve learning objects (see section 22.1). These repositories should enable their different users to retrieve and select *appropriate* learning materials – which is the classic context for adaptation. However, successful approaches must include open corpus adaptive hypermedia, for example, becoming flexible enough to deal with varying metadata schemes and metadata details / quality.

Possible solutions discussed in section 22.3.2 of this paper include the manual indexing approach, which links hypermedia documents to external models and the automatic keyword-based indexing approach. Most promising in this context is on-the-fly metadata identification approach also presented in section 22.3.2. By analyzing usage patterns and signature structures in large hypertexts, metadata-like information can be gathered and explored to show relations between documents, rank documents or relations, and recommend relevant documents or relations for specific target groups.

22.4.2 Interoperability and the Open Corpus Problem

Apart from the re-use of content, which might be the most obvious implication of the open corpus problem, the *re-use of adaptive functionality* itself can be seen as equally important. Currently, most adaptive hypermedia systems are built from scratch, re-implementing adaptive functionality instead of re-using appropriate software modules. A first step to arrive at a re-usable adaptive functionality is to analyze and describe adaptive functionality in a system-independent manner, which, formally stated, describes the adaptation algorithms together with the required processing data. This processing data pertains to all aspects of the adaptation process: the adaptation-specific information in the adaptive hypermedia system, the user characteristics and models, as well as data that is only available at runtime (e.g., [31], which introduces a formal characterization of adaptive functionality in some of the most-cited adaptive educational hypermedia systems).

The re-use of adaptive functionality across applications requires interoperability solutions for adaptive systems. Interoperability is a very important aspect of today's systems, not only adaptive systems, and many issues for enabling true interoperability remain to be solved.

In section 22.3.2, we have seen an approach for solving the open corpus problem on the level of architectures. A service-oriented architecture with personalization services – each of them realizing a certain adaptive functionality – is proposed. Integration and syndication of the results of the services is realized within a dedicated reasoning component, making this reasoning a very important part of the inter-operation process.

We claim that solutions to the open corpus problem in adaptive hypermedia contribute to solving general interoperability issues, and on the other hand, interoperable adaptive hypermedia systems have—in one way or the other—have tackled and continue to contribute solutions to the open corpus problem. Furthermore, continuous efforts are required to solve re-usability of adaptive functionality and adaptive systems, and interoperability between adaptive components or systems. As of today, adaptive hypermedia systems are mainly developed at universities, with limited commercial use. While evaluations of adaptive hypermedia systems have proven their benefit, the wide use of these methods and techniques in practical systems is still pending. One of the reasons for this arises from the missing or limited re-usability. Development costs are high, since in the majority of cases the realization of adaptive hypermedia techniques starts from scratch instead of extending and re-using existing systems. Re-use can help in limiting development costs, and lower development costs will make it more attractive for developers and project managers to choose adaptive, personalized solutions.

22.4.3 Adaptive Hypermedia and the Semantic Web

Overcoming the open corpus problem in adaptive hypermedia receives special importance in the light of upcoming expectations and research on adding semantics to Web information [3]. The need for personalized, adaptive access to Web information will be high if semantic-enabled applications want to demonstrate their effectiveness: one-size-fits-all approaches will not explore the full potential that can be found with automated reasoning that is based on machine-processable semantics. On the other hand, the information space of the (Semantic) Web can be characterized as highly dynamic, open, and heterogeneous: Far from being under control of only a couple of system developers, information on the Web can emerge, be modified, altered, or disappear. User-tailored applications in the (Semantic) Web therefore require open-world solutions.

The Semantic Web (see also Chapter 23 of this book [26]) aims at machine-readable and machine-processable semantics for the Web. Metadata, together with formal ontologies providing the semantics, are a meaningful source for expressing adaptation-specific information.

As an example, the document-to-concept relations discussed in section 2.3 can be expressed in the language of the Resource Description Framework (RDF [62]), with direct references to an ontology (e.g., written in the language OWL [61]). The ontology itself can be used for expressing the required adaptation-specific information in complementary models. The crucial point in document-to-concept relations, the intrinsic dependency on specific concept models, can be tackled by using different ontologies corresponding to the different concept models, and applying techniques from ontology mapping and ontology merging to externalize this intrinsic information. Furthermore, the languages of the Semantic Web provide the required add-on to pure metadata approaches: For example, by analyzing whether adaptation-specific information can be encoded with standard metadata catalogs for learning materials, constraints on subject classifications can be identified. The embedding of the subject classifier in a concept model can be described with the languages on the Semantic Web in a machine-processable way, thus enabling adaptation algorithms to evaluate and reason about the subject classifier and its meaning with respect to a referenced ontology.

The layered architecture of the Semantic Web, accompanied by reasoning engines, rule languages, logical formalisms, and trust models, provides means for reasoning about the adaptation-specific information in a standardized but open environment. This facilitates, on the one hand, an adaptation functionality that processes this information in order to determine appropriate adaptive treatment. On the other hand, the reasoning can—at least to some degree—be proven and externalized to explain to end-users what has been done. This will improve the transparency of the whole adaptation process as users can inquire about why a certain recommendation or navigation support or whatever adaptive treatment has been determined by the adaptation module. In addition, various possibilities thus open up for extending controllability of the adaptation process, leading to *scrutable* adaptation (see the discussion on scrutable user models in [39]).

Overall, we can observe that the Semantic Web, with its languages, formalisms, and machine-processable semantics, provides excellent conditions for the use of adaptation. The Semantic Web achieves the separation of Web content from its later delivery and a certain context-of-use by enabling computer programs to reason about this Web content and its meaning. Adaptation, on the other hand, allows for the tailoring of this delivery according to the specific and individual requirements of users, within their current context. In this way, adaptation or personalization is important for optimizing the process of querying for, retrieving, selecting, and accessing information on the Web under user-specific constraints, and adaptive methods from the field of Adaptive Hypermedia should be realized very well within this Semantic Web architecture.

22.5 Conclusions

Since the mid-nineties, techniques in the field of adaptive hypermedia have been developed to adapt hypertexts to the needs of individual users. Success stories of adaptive hypermedia have especially been reported in the educational field, with the delivery of different individual learning paths and recommendations for learning goals or exercises, thus providing precisely attuned guidance and support during the learning process.

Despite the fact that techniques from adaptive hypermedia have proven their successfulness in providing individually optimized views on large hypertextual information spaces, wide-spread use of these techniques in e-learning is still pending. We argue that one reason for this can be identified as the open corpus problem in adaptive hypermedia.

This paper provides an in-depth analysis of the open corpus problem in adaptive hypermedia within the application area of educational systems. We show how document corpora and adaptation techniques are intertwined, and discuss consequences or this coupling for applying adaptive hypermedia techniques to open and dynamic information spaces. We characterize and compare the different approaches to overcome the open corpus problem and discuss their benefits and drawbacks. We reveal the relations of the open corpus problem to the re-usability and interoperability of adaptive systems, and point out the benefits of applying Semantic Web technologies to tackle and solve the open corpus problem.

Acknowledgments. This research is supported by the National Science Foundation under Grant No. 0447083 and partially supported by REWERSE - Reasoning on the Web (www.rewerse.net), Network of Excellence, 6th European Framework Program.

References

1. Allan, J.: Automatic hypertext link typing. In: Proc. of 7th ACM Conference on Hypertext. ACM Press (1996) 42-52
2. Bareiss, R., Osgood, R.: Applying AI models to the design of exploratory hypermedia systems. In: Proc. of Fifth ACM Conference on Hypertext. ACM Press (1993) 94-105

3. Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web. Scientific American, May (2001) 35-43
4. Billsus, D., Pazzani, M.: Adaptive news access. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.): The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321. Springer-Verlag, Berlin Heidelberg New York (2007) this volume
5. Bollen, J., Heylighen, F.: A system to restructure hypertext networks into valid user models. The New Review of Multimedia and Hypermedia 4 (1998) 189-213
6. Brin, S., Page, L.: The anatomy of a large-scale hypertextual (Web) search engine. In: Ashman, H., Thistewaite, P. (eds.) Proc. of Seventh International World Wide Web Conference. Vol. 30. Elsevier Science B. V. (1998) 107-117
7. Brusilovsky, P.: Adaptive hypermedia. User Modeling and User Adapted Interaction 11, 1/2 (2001) 87-110
8. Brusilovsky, P.: Adaptive navigation support. In: Brusilovsky, P., Kobsa, A., Neidl, W. (eds.): The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321. Springer-Verlag, Berlin Heidelberg New York (2007) this volume
9. Brusilovsky, P., Chavan, G., Farzan, R.: Social adaptive navigation support for open corpus electronic textbooks. In: De Bra, P., Nejdl, W. (eds.) Proc. of Third International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH'2004). Lecture Notes in Computer Science, Vol. 3137. Springer-Verlag (2004) 24-33
10. Brusilovsky, P., Eklund, J.: A study of user-model based link annotation in educational hypermedia. Journal of Universal Computer Science 4, 4 (1998) 429-448
11. Brusilovsky, P., Eklund, J., Schwarz, E.: Web-based education for all: A tool for developing adaptive courseware. In: Ashman, H., Thistewaite, P. (eds.) Proc. of Seventh International World Wide Web Conference. Vol. 30. Elsevier Science B. V. (1998) 291-300
12. Brusilovsky, P., Millán, E.: User models for adaptive hypermedia and adaptive educational systems. In: Brusilovsky, P., Kobsa, A., Neidl, W. (eds.): The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321. Springer-Verlag, Berlin Heidelberg New York (2007) this volume
13. Brusilovsky, P., Pesin, L.: Adaptive navigation support in educational hypermedia: An evaluation of the ISIS-Tutor. Journal of Computing and Information Technology 6, 1 (1998) 27-38
14. Brusilovsky, P., Rizzo, R.: Map-based horizontal navigation in educational hypertext. Journal of Digital Information 3, 1 (2002) http://jodi.ecs.soton.ac.uk/Articles/v03/i01/Brusilovsky/
15. Brusilovsky, P., Schwarz, E.: Concept-based navigation in educational hypermedia and its implementation on WWW. In: Müldner, T., Reeves, T.C. (eds.) Proc. of ED-MEDIA/ED-TELECOM'97 - World Conference on Educational Multimedia/Hypermedia and World Conference on Educational Telecommunications. AACE (1997) 112-117
16. Burke, R.: Hybrid Web recommender systems. In: Brusilovsky, P., Kobsa, A., Neidl, W. (eds.): The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321. Springer-Verlag, Berlin Heidelberg New York (2007) this volume
17. Campione, M., Walrath, K.: The Java Tutorial. Addision-Wesley, (2003)
18. Carmona, C., Bueno, D., Guzmán, E., Conejo, R.: SIGUE: Making Web Courses Adaptive. In: De Bra, P., Brusilovsky, P., Conejo, R. (eds.) Proc. of Second International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH'2002). Lecture Notes in Computer Science, Vol. 2347. Springer-Verlag (2002) 376-379
19. Carr, L., Hall, W., Bechhofer, S., Goble, C.: Conceptual linking: Ontology-based open hypermedia. In: Proc. of 10th International World Wide Web Conference. ACM Press (2001) 334-342

20. Chen, H., Houston, Andrea L., Sewell, Robin R., Schatz, Bruce R.: Internet browsing and searching: User evaluations of category map and concept space techniques. Journal of the American Society for Information Science 49, 7 (1998) 582-608
21. Cleary, C., Bareiss, R.: Practical methods for automatically generating typed links. In: Proc. of 7th ACM Conference on Hypertext. ACM (1996) 31-41
22. De Bra, P., Calvi, L.: AHA! An open Adaptive Hypermedia Architecture. The New Review of Hypermedia and Multimedia 4 (1998) 115-139
23. Dieberger, A.: Where did all the people go? A collaborative Web space with social navigation information (2000) available online at http://homepage.mac.com/juggle5/WORK/publications/SwikiWriteup.html
24. Dieberger, A., Guzdial, M.: CoWeb - experiences with collaborative Web spaces. In: Lueg, C., Fisher, D. (eds.): From Usenet to CoWebs: Interacting with Social Information Spaces. Springer-Verlag, New York (2003) 155-166
25. Dolog, P., Henze, N., Neidl, W.: Logic-based open hypermedia for the semantic web. In: Proc. of 1st International Workshop on Hypermedia and the Semantic Web at Hypertext 2003. (2003) http://www.l3s.de/~dolog/pub/htsw2003.pdf
26. Dolog, P., Nejdl, W.: Semantic Web Technologies for the Adaptive Web. In: Brusilovsky, P., Kobsa, A., Neidl, W. (eds.): The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321. Springer-Verlag, Berlin Heidelberg New York (2007) this volume
27. Gauch, S., Speretta, M., Chandramouli, A., Micarelli, A.: User profiles for personalized information access. In: Brusilovsky, P., Kobsa, A., Neidl, W. (eds.): The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321. Springer-Verlag, Berlin Heidelberg New York (2007) this volume
28. Goy, A., Ardissono, L., Petrone, G.: Personalization in e-commerce applications. In: Brusilovsky, P., Kobsa, A., Neidl, W. (eds.): The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321. Springer-Verlag, Berlin Heidelberg New York (2007) this volume
29. Green, S.: Building hypertext links by computing semantic similarity. IEEE Transactions on Knowledge and Data Engineering 11, 5 (1999) 713-730
30. Henze, N.: Personal Readers: Personalized Learning Object Readers for the Semantic Web. In: Looi, C.-K., McCalla, G., Bredeweg, B., Breuker, J. (eds.) Proc. of 12th International Conference on Artificial Intelligence in Education, AIED'2005. IOS Press (2005) 274-281
31. Henze, N., Neidl, W.: A Logical Characterization of Adaptive Educational Hypermedia. New Review of Hypermedia and Multimedia 10, 1 (2004) 77-113
32. Henze, N., Nejdl, W.: Adaptivity in the KBS Hyperbook System. In: Brusilovsky, P., Bra, P.D., Kobsa, A. (eds.) Proc. of Second Workshop on Adaptive Systems and User Modeling on the World Wide Web. Vol. 99-07. Eindhoven University of Technology (1999) 67-74, also available as http://wwwis.win.tue.nl/asum99/henze/henze.html
33. Henze, N., Nejdl, W.: Extendible adaptive hypermedia courseware: Integrating different courses and Web material. In: Brusilovsky, P., Stock, O., Strapparava, C. (eds.) Proc. of Adaptive Hypermedia and Adaptive Web-based Systems, AH'2000. Lecture Notes in Computer Science, Vol. 1892. Springer-Verlag (2000) 109-120
34. Henze, N., Nejdl, W.: Adaptation in open corpus hypermedia. International Journal of Artificial Intelligence in Education 12, 4 (2001) 325-350
35. Henze, N., Nejdl, W.: Knowledge modeling for open adaptive hypermedia. In: De Bra, P., Brusilovsky, P., Conejo, R. (eds.) Proc. of Second International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH'2002). Lecture Notes in Computer Science, Vol. 2347. (2002) 174-183
36. Hirashima, T., Hachiya, K., Kashihara, A., Toyoda, J.i.: Information filtering using user's context on browsing in hypertext. User Modeling and User Adapted Interaction 7, 4 (1997) 239-256

37. IEEE LTCS WG2: PAPI Learner, Draft 8 Specification. IEEE P1484.2 Learner Model Working Group (2002) available online at http://edutool.com/papi/
38. IEEE LTCS WG12: IEEE Standard for Learning Object Metadata, IEEE (2002) available online at http://ltsc.ieee.org/wg12/par1484-12-1.html
39. Kay, J., Kummerfeld, B.: Scrutability, user control and privacy for distributed personalization. In: Proc. of Workshop on Privacy-Enhanced Personalization at CHI 2006. (2006) http://www.isr.uci.edu/pep06/papers/PEP06_KayKummerfeld.pdf
40. Kibby, M.R., Hayes, J.T.: Towards intelligent hypertext. In: McAleese, R. (ed.) Hypertext: Theory into practice. Intellect, Oxford (1989) 164-171
41. Kohonen, T.: Self-organizing maps. Springer Verlag, Berlin (1995)
42. Kohonen, T.: Self-organization of very large document collections: State of the art. In: Niklasson, L., Bodén, M., Ziemke, T. (eds.) Proc. of 8th International Conference on Artificial Neural Networks, ICANN98. Vol. 1. Springer (1998) 65-74
43. Lagoze, C., al., e.: Core services in the architecture of the national science digital library (NSDL). In: Marchionini, G., Hersh, W. (eds.) Proc. of ACM/IEEE-CS Joint Conference on Digital Libraries. ACM Press (2002) 201-209
44. Macedo, A.A., Pimentel, M.d.G.C., Camacho-Guerrero, J.A.: An infrastructture for open latent semantic linking. In: Anderson, K.M., Moulthrop, S., Blustein, J. (eds.) Proc. of 13th ACM Conference on Hypertext and Hypermedia (Hypertext 2002). ACM (2002) 107-115
45. Masthoff, J.: Automatic generation of a navigation structure for adaptive Web-based instruction. In: Brusilovsky, P., Henze, N., Millán, E. (eds.) Proc. of Workshop on Adaptive Systems for Web-Based Education at the 2nd International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH'2002). (2002) 81-91, also available at http://www.csd.abdn.ac.uk/~jmasthof/ah02workshop.pdf
46. Mayes, J.T., Kibby, M.R., Watson, H.: StrathTutor: The development and evaluation of a learning-by-browsing on the Macintosh. Computers and Education 12, 1 (1988) 221-229
47. Micarelli, A., Gasparetti, F.: Adaptive focused crawling. In: Brusilovsky, P., Kobsa, A., Neidl, W. (eds.): The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321. Springer-Verlag, Berlin Heidelberg New York (2007) this volume
48. Micarelli, A., Gasparetti, F., Sciarrone, F., Gauch, S.: Personalized search on the World Wide Web. In: Brusilovsky, P., Kobsa, A., Neidl, W. (eds.): The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321. Springer-Verlag, Berlin Heidelberg New York (2007) this volume
49. Micarelli, A., Sciarrone, F., Marinilli, M.: Web document modeling. In: Brusilovsky, P., Kobsa, A., Neidl, W. (eds.): The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321. Springer-Verlag, Berlin Heidelberg New York (2007) this volume
50. Millard, D.E., Moreau, L., Davis, H.C., Reich, S.: FOHM: a fundamental open hypertext model for investigating interoperability between hypertext domains. In: Proc. of Eleventh ACM on Hypertext and hypermedia. ACM Press (2000) 93-102
51. Murray, T.: MetaLinks: Authoring and affordances for conceptual and narrative flow in adaptive hyperbooks. International Journal of Artificial Intelligence in Education 13, 2-4 (2003) 199-233
52. Pazzani, M.J., Billsus, D.: Content-based recommendation systems. In: Brusilovsky, P., Kobsa, A., Neidl, W. (eds.): The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321. Springer-Verlag, Berlin Heidelberg New York (2007) this volume
53. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgen Kaufmann Publishers, Inc., San Mateo, CA (1988)

54. Rizzo, R.: A Method of Labelling Hypertext Links in a Context - Dependent Way. In: Davies, G., Owen, C. (eds.) Proc. of WebNet'2000, World Conference of the WWW and Internet. AACE (2000) 449-453
55. Rizzo, R., Allegra, M., Fulantelli, G.: Developing hypertexts through a self-organized map. In: Maurer, H., Olson, R.G. (eds.) Proc. of WebNet'98, World Conference of the WWW, Internet, and Intranet. AACE (1998) 768-773
56. Schafer, J.B., Frankowski, D., Herlocker, J., Sen, S.: Collaborative filtering recommender systems. In: Brusilovsky, P., Kobsa, A., Neidl, W. (eds.): The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321. Springer-Verlag, Berlin Heidelberg New York (2007) this volume
57. Shipman III, F.M., Marshall, C.C.: Spatial hypertext: an alternative to navigational and semantic links. ACM Computing Surveys 31, 4es (1999) Article No. 14
58. Smith, A.S.G., Blandford, A.: MLTutor: An application of machine learning algorithms for an adaptive Web-based information system. International Journal of Artificial Intelligence in Education 13, 2-4 (2003) 235-261
59. Smyth, B.: Case-base recommendation. In: Brusilovsky, P., Kobsa, A., Neidl, W. (eds.): The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321. Springer-Verlag, Berlin Heidelberg New York (2007) this volume
60. Tudhope, D., Taylor, C.: Navigation via similarity: automatic linking based on semantic closeness. Information Processing & Management 33, 2 (1997) 233-242
61. W3C: OWL: Web Ontology Language Reference. World Wide Web Consortium (2004) available online at http://www.w3.org/TR/owl-ref/
62. W3C: Resource Description Framework (RDF): Concepts and Abstract Syntax. World Wide Web Consortium (2004) available online at http://www.w3.org/TR/owl-ref/
63. W3C: WSDL: Web Services Description Language, version 2.0. World Wide Web Consortium (2004) available online at http://www.w3.org/TR/2004/WD-wsdl20-20040803/
64. Weber, G., Brusilovsky, P.: ELM-ART: An adaptive versatile system for Web-based instruction. International Journal of Artificial Intelligence in Education 12, 4 (2001) 351-384
65. Weber, G., Specht, M.: User modeling and adaptive navigation support in WWW-based tutoring systems. In: Jameson, A., Paris, C., Tasso, C. (eds.) Proc. of 6th International Conference on User Modeling. SpringerWienNewYork (1997) 289-300
66. Yan, T.W., Jacobsen, M., Garcia-Molina, H., Dayal, U.: From user access patterns to dynamic hypertext linking. Computer Networks and ISDN Systems. (1996) 1007-1014
67. Yee, K.-P., Swearingen, K., Li, K., Hearst, M.: Faceted metadata for image search and browsing. In: Proc. of ACM Conference on Human Factors in Computing Systems, CHI 2003. ACM Press (2003) 401-408