# Enhancing the Adaptivity of an Existing Website with an Epiphyte Recommender System

## Bruno Richard, Pierre Tchounikine

# Enhancing the Adaptivity of an Existing Website with an Epiphyte Recommender System

Bruno Richard, Pierre Tchounikine

LIUM (Laboratoire d'Informatique de l'Université du Maine)
FRE CNRS 2730
Avenue Laennec, 72085 Le Mans Cedex 9, France
Fax : +33 2 43 83 38 68
{bruno.richard, pierre.tchounikine}@lium.univ- lemans.fr

## Abstract

In this paper we propose an approach to enhance the adaptivity of an existing Website by plugging on top of it ("epiphyte approach") a recommender system that displays additional tips and functionalities in a separate window. The recommender system analyzes the way the user browses through the Website according to predefined prototypical ways of using the Website ("models of use") and then proposes information or functionalities that appear useful according to this model of use. Different models of use can be identified, each of them corresponding to a "logical extension" of the original Website. Associating an existing Website with such logical extensions therefore allows enhancing its adaptivity whilst (1) not modifying the original Website and (2) facilitating the evolution of the adaptive features as this only requires modifying the recommender system. This approach can be used as an alternative and/or in association with other approaches related in the literature.

## 1. INTRODUCTION

Analyzing logs and/or users' interviews often highlights that the way in which a Website is effectively used is slightly different from the Website promoters' and designers' expectations: pages or entire parts of Websites that remain unused, functionalities offered by the Website that are not used or in a different manner than the one that was expected, users that develop idiosyncratic means to address objectives that were not anticipated by the Website designers, etc. This is particularly the case for complex Websites where the interactions with users cannot be reduced to an initial request, i.e., the user has to navigate and achieve different actions to identify what he is looking for and/or discover, on the way, some interesting material he was not aware of. Although this problem is always the result of different interrelated issues, one of these is that the design and maintenance of Websites must face an increasing problem: addressing the satisfaction of different types of users, i.e., users that can address different goals and/or present different skills or

different ways to navigate. Classical Websites that only address an "ideal" unique user and present the same behaviour independently from the user's goal, his experience or his navigation through the Website structure are too static and cannot match a diversity of expectations. Complex Websites must be able to adapt themselves to their users, either to the idiosyncratic behaviour of individuals, or to different prototypical ways of using the Website.

Addressing the adaptivity of a Website is a problem that must face both primo-design and emergence issues. From a general point of view, the democratization of the Web leads more and more different users to employ the Web, different users that present different profiles. The fact that the Website addresses an open community or an identified community and the characteristics of this community (or communities) and of its members must be carefully studied. However, in both cases, a central issue is that how a Website is used in concrete terms emerges and evolves during time and cannot be completely anticipated when the Website is first designed.

In this paper we propose an approach to enhance the adaptivity of an existing Website with an epiphyte recommender system. The addressed problem is that of an existing complex Website from which an analysis of the effective use has highlighted that it does not match the users' or promoters' expectations, i.e., a Website whose structure is not well adapted for certain uses. Modifying or rebuilding such a Website is generally too expensive and in any case a short-view approach for the two reasons we have highlighted previously: (1) the design of a Website must consider different ways of using them and not one single ideal use and (2) the way users use a Website evolves. We therefore propose to keep the Website unchanged and to plug a recommender system on top of it ("epiphyte approach"). A recommender system is a system that gives the Website users tips and functionalities that are displayed in an additional window that comes on the top of the Website window. A user browsing the Website is therefore presented with (1) the Website response and (2) the recommender system's additional information and functionalities. These additional information, tips and/or functionalities are generated on the fly from an analysis of the user's navigation. The recommender system therefore acts as an "adaptive extension" of the original Website.

In this work we consider prototypical uses of the Website that we call "models of use" (by opposition to considering idiosyncratic behaviours of individuals, i.e., attempting to construct individual profiles of users). The recommender system analyzes the way the user browses through the Website according to a given model of use and then proposes the appropriate tips or functionalities that appear useful (according to this model) for the user. Different models of use can be identified, each of them corresponding to a "logical extension" of the original Website. Associating an existing Website with such logical extensions therefore allows to enhance its adaptivity whilst (1) not modifying the original Website and (2) facilitating the overall maintenance (e.g., taking into account of the emergence of new ways of using the Website or modification of the Website) as the corresponding

modifications only require the creation or modification of a model of use (and eventually its associated functionalities) within the recommender system.

The paper is organized as follows. In section 2 we describe our general concern, present a quick overview of adaptive hypermedia techniques (that can be used to address this issue) and their limits, and then introduce our approach. In section 3 we present in detail the putting into practice of this approach. In section 4 we describe the system architecture and implementation issues. In Section 5 we propose a discussion of different aspects of the proposed approach and Section 6 summarizes the work.

The examples that are used in this paper are issued from the application of the proposed approach to the design of a recommender system prototype for a Website created by the French Ministry of Education and Research, in the context of an industrial contract that has partially supported this work. Although this recommender system was finally not open to the public (for extra-scientific reasons), its development allowed us to validate our approach and the system implementation.

## 2. INTRODUCING ADAPTIVITY IN COMPLEX WEBSITES

### 2.1. General concern

In this work we consider complex Websites and - as an example - complex repository-like Websites. Within such a Website, users define a criterion and/or select a topic from a list and/or type a keyword, the system searches the items that correspond in the underlying database and displays the results, generating a page on the fly from the user's request or navigation. In some cases a unique research interface is sufficient and the Website remains simple. However, in some cases, the structure of the Website is made complex by the fact that it proposes several levels (e.g., general orientation means at the top-level page that orientates towards different sublevel interfaces), different means (e.g., simple keywords search, multiple keyword search, criteria editor or taxonomy-based search) that can vary according to the Website levels, different functionalities (e.g., selecting items, organising data, restricting data, expanding data or connecting to other related Websites) according to the current state of the search and/or the browsing path, etc.

When constructing such Websites designers have to face the problem of the interface usability. Generally, the interface is designed to be understood by a large panel of users, targeting an "ideal" standard user. The classical problem (not limited to Websites) of the design of such systems is that an interface designed for a "standard" user often appears inadequate for a "novice" user or for an "expert" user: while the first one can get lost in the Website structure or in functionalities too complicated, the second one can miss interesting functionalities or pages not clearly announced to avoid novices' cognitive overload. With repository-like complex Websites the problem is not only of "novice" and "expert" users, but of users that can be very heterogeneous and that can have different aims (e.g., browsing through

the Website VS searching a well identified item, being interested in different characteristics of the repository items, etc.). A static interface forces the designers to compile all the different functionalities in the same screens and to propose these screens independently of the users' aims or skills, when it is known that different users require different means (for instance, keyword-search is often inadequate for basic users (1)) and that proposing a high-number of functionalities appears as disturbing rather than useful.

Solutions to these problems can be found in Adaptive Hypermedia techniques.

## 2.2. Adaptive Hypermedia techniques

Techniques developed in Adaptive Hypermedia can be used in different ways to address Website adaptivity requirements, e.g.:

- Adaptation of the pages that are proposed to the user, using user modelling techniques (construction and maintenance of a user model) (2), (3).

- Recommendations, e.g., indication of pages that could be of interest for the users. As an example, "link augmentation" is a technique that allows inserting dynamically additional links into existing Webpages (4). Different approaches have been proposed, using techniques such as user modelling, concepts extraction (contend based techniques, using metadata and ontologies) (5), (6) or collaborative-based mechanisms (comparisons between pages or users) (7).

- Guidance, e.g. presenting users with the « spatial context » (information on the domain and the links between its different elements), « preview cues » (links that could be of interest) or « dimensional sorting » (sorting data according to predefined keys) (1).

These techniques, that have been originally more particularly developed in the context of Adaptive Educational Hypermedia (8), have proved to be powerful, and adapting them to the Web context is a hot topic (see for example (9)). However, they present a certain number of drawbacks that can rule them out in certain contexts:

- Integration of the adaptive issues in the Website design. Websites using Adaptive Hypermedia techniques allow the generation of pages on the fly according to the users' needs and thus "design decisions need to be specified beforehand" (10). It appears difficult to apply such techniques on an existing Website without re-engineering it.

- Necessity for an ontology. Many techniques are based on an ontology of the concerned domain (typically, use of an ontology to classify natural language terms and then propose links (11)). However, such an ontology is not always available or can be too expensive to build and maintain.

- "Cold start" problem. Many techniques are based on the comparison of users' profiles (7) or of pages (using for instance the K-Nearest Neighbours algorithm (12)). However these different techniques suffer from the "cold start" problem: the system is only able to propose suitable links after different pages have been browsed and many users' profiles built. To avoid such data scarcity systems can interact directly with users (e.g., in (13) the system allows users to indicate if a recommendation is interesting or not). However, such techniques cannot be used in certain cases, for instance in contexts where there is a risk that most users may skip to another Website instead of answering questions.

- Necessity for metadata extracted from Web pages (14) or defined by the Website designers (6). Automatic extraction techniques require ontologies, which as explained above are not always available. Defining and maintaining metadata for each Website page is a tedious task that can become repulsive and/or intractable for complex Websites.

## 2.3. Introducing adaptivity with a recommender system

The approach we propose is an alternative and/or complementary approach to current works. It can be defined as a combination of task-based performance support and site-based recommender services. Its basic principles are as follows.

1. Non-modification of the original Website. The fact that a Website is easy to modify and improve should be a top-level issue of the design. However, this is not often the case. Modifying or rebuilding existing Websites is generally too expensive. The objective of our work is therefore to enable the enhancement of the adaptivity of an existing Website without modifying it. Another argument in favour of this approach is that the structure of a Website and the way users use a Website evolves. A "logical" approach, i.e., an approach that allows introducing adaptivity features and then adapting these features whilst not modifying the original Website is therefore interesting.

2. Epiphyte approach. The approach we propose in order to introduce adaptivity features whilst not modifying the original Website is to plug on top of the original Website a system that introduces the required functionalities. This approach is known as the "epiphyte approach" (an epiphyte system is a system that is associated with another without modifying it (15), (16), (17)). Such a system comes as a frontal (from an architectural point of view: a proxy) to the original Website. This allows meeting the central requirement (non-modification of the original Website) and presents different advantages (see *infra*). One of these is that the epiphyte system can be associated to a Website whilst being physically distant (on another server, in another site). Such a loose association introduces

a very interesting flexibility. Another advantage is that the overall architecture of the epiphyte system is generic and can be used for different Websites.

3. Recommender system. Our approach consists in introducing adaptivity issues by presenting users with the original Website features and additional features (information, tips and functionalities), by opposition to the adaptation of the original Website features. Our system is a recommender system whose output is displayed in an additional window that comes on top of the original Website window. This additional information, tips and/or functionalities are generated on the fly from an analysis of the user's navigation. The recommender system therefore acts as an "adaptive extension" of the original Website. It can be deactivated at any moment.

4. Task-based performance. Because of the cold-start problem (and because we know the limits of user modelling techniques in the context of open communities) we propose an approach that is not based on user modelling, i.e. capturing the individual characteristics of a user. Our approach is based on the modelling of the Website structure and the identification of prototypical uses of the studied Website, that we call "models of use". A model of use denotes a prototypical way to achieve a prototypical task that a user can intend to achieve through the considered Website (e.g., for a repository-like Website: "browse for information", "search a particular item", "collect items", etc.). These models are used as references by the recommender system in order to analyze the user's navigation and present tips or functionalities that appear useful according to this model.

The approach we propose is therefore model-based (modelling the Website structure and the models of use are the key issues). It does not require the construction of a domain-ontology or metadata annotation of the Website content as it is based on the analysis of the navigation in respect to a given model and not on the items that the Website contains. It can however be enhanced by adaptive functionalities based on such an ontology. Similarly, individualization issues can be added by integrating user modelling techniques to our approach.

# 3. BUILDING A RECOMMENDER SYSTEM TO ENHANCE EXISTING WEBSITE ADAPTIVITY

## 3.1. Definitions

Our approach is based on the use of different models. These models are described as graphs. In these graphs, states represent Website pages (or types

of Website pages if generic URLs are considered) and arcs represent links between pages.

**Website graph:** The Website graph corresponds to the set of pages and links proposed by a given Website.

**Graph of the recommender system:** The graph of the recommender system corresponds to the set of pages and links that are considered by the recommender system. This includes pages and links of the considered Website (i.e., the Website graph) and the pages and links that are generated by the recommender system. The Website that the user gets corresponds to the graph of the recommender system. The pages and links of this graph are associated with the additional tips and functionalities that will be proposed by the recommender system.

**Model of use:** A model of use is a subset of the graph of the recommender system that denotes a prototypical way of using the Website to achieve a given identified task.

**Trace:** A trace is a subset of the graph of the recommender system that denotes the pages and links browsed within a session by a user.

Modelling a typical way of using a hypermedia has been used in some systems such as Hynecosum (18). This system is based on the definition of typical categories of users, each of them using the hypermedia to achieve a different given task. This is very close to our notion of models of use, that can be viewed as a static user model which is a stereotype or, in other words, a "group model". Differently from works as (19) we do not attempt to construct and maintain individual models to propose individualized issues; this is however a possible line for future research.

### 3.2. General method

We have implemented a generic architecture that proposes the basic functionalities of our epiphyte approach (cf. section 4). Building a recommender system by instantiating this architecture requires the following stages:

1. Analysis of the Website and of its effective uses and definition of the general requirements. The structural analysis of the Website allows the construction of the Website graph. The Website graph and data denoting the way the Website is effectively used (different sources can be used, e.g. traces constructed from log-analysis, users' interviews) are the basis for fixing the issues to be addressed (ergonomics problem, lack of explanation, unused functionality, emergent unexpected use, etc.).

2. Definition of the models of use (i.e., the prototypical ways of using the Website that one wants to support) and the associated tips and the functionalities. The overall process conducts to the

construction of the graph of the recommender system, which is the basis for defining the help that is proposed to the user.

3. Implementation of these additional functionalities.



**Figure 1: Home page of the Educasource Website**

In the following subsections we describe these different phases with examples from the building of such a recommender system for the Educasource Website, a Website constructed by the French Ministry of Education and Research. Educasource is (in fact, was: the Website has been recently deactivated) a repository dedicated to teachers. It proposes educational resource descriptors (different items describing an educational resource such as the author, the domain, the curriculum level, the type of resource - Website, CD-Rom, etc. -, a summary, keywords, etc.). Educasource structure is prototypical of repository Websites. It proposes two main research functionalities, a basic keyword search engine (an advanced search is also proposed) and a hierarchic taxonomy (in this case Dewey's taxonomy). A forum is proposed to allow teachers using the Website to communicate and exchange interesting pages or links. Figure 1 presents the home page of the Educasource Website. From this front-page users can type a keyword and will be presented with a list of resource descriptors, select one of the taxonomy items and then (recursively) one of the sub-items until they reach the leaves of the taxonomy and the corresponding list of resource descriptors or access the Forum.

### 3.3. Analysis of the Website and of its effective uses and definition of the general requirements

The first phase to be achieved is the construction of the Website graph and then a diagnostic of how the Website is used in concrete terms. In the case of Educasource, the effective use of the Website and the users' behaviours was analysed one year after it was put on-line (accessible to the open community of teachers) by a private company (external audit). This analysis (and the one we did when building the recommender system) highlighted a certain number of differences between the effective use of the Website and the Website promoters' expectations. As examples:

- Unused functionalities. For instance, the forum was totally ignored whereas it was an expectation of the Website promoters that the forum would be used by teachers to exchange experiences or share interesting pages and would help the Website to appear dynamic.

- Unbalanced uses of complementary functionalities. It appeared that the keyword search engine was principally used, to the detriment of the taxonomy search method.

- Different surface problems (e.g., the lack of some functionalities such as data sorting or data restructuration at some level of the Website or disorientation problems).

The fact that a Website is used differently from the promoters' expectations is not necessarily a problem. The examples we have outlined here-above are however problematical from different points of view. For instance, the two search functionalities (keyword and taxonomy) are proposed because they are complementary. The use of keyword-search (respectively, taxonomy-search) allows accessing resources one would never be aware of with the other method. Another more general problem linked to this pre-eminence of the keyword search engine is that although it appears intuitively simple, keyword search is often inadequate for basic users (1), in particular when they have to deal with too long or too short lists of heterogeneous results (an answer composed of more than 20 resource descriptors becomes intractable). As most users are not skilled Web-users and/or do not have very good working conditions (difficult access to computers, Modem lines) and/or are just gathering to see if something would be of interest, keyword search difficulties and limitations are a factor of the Website under-exploitation. The fact that the Forum remains unused is also considered as disappointing and possibly discouraging for dynamic users.

If one considers things from a more general level, for its promoters, such a Website is an instrument that aims at a certain policy, in this case promote new pedagogical practices. As an example, there was an expectation that teachers would exchange experiences one from another through the Forum or would take benefit from the existence of other related Websites that address

pedagogical issues (official curricula, etc.). The diagnosis highlighted that it was not the case; the emergent uses do not meet the promoters' expectations.

This example highlights that the design of a Website often corresponds to objectives and/or principles that are not apparent and that when one attempts to merge different requirements within a single interface the result can be inadequate. The Website structure is "neutral"; it allows different ways of using it but specifically supports none of them. Our approach consists in expliciting the different ways of using the Website that one wants to promote and to present the users with "logical extensions" of the Website (by the means of an additional window) that are specifically adapted to these alternatives.

Identifying these different ways of using the Website is a problem in itself. It can be addressed by studying the Website logs and attempting to automatically define users' behaviour stereotypes. In this work, given our context and, in particular, the objective of promoting identified behaviours, we adopted an approach based on the *a priori* modelling.

## 3.4. Definition of the models of use and the associated functionalities

### 3.4.1. Definition of the models of use

A "Model of use" denotes a prototypical way of using the Website to achieve a given predefined task. Such a model can correspond to (1) a particular use of the Website that emerges (i.e., one establishes that some users recurrently behave this way) and that one wants to support or (2) to a prescriptive model that one wants to promote for commercial, communication or pedagogical issues (the objective is then to influence the users' navigation and make the users browse some particular pages or use some particular functionalities).

**Figure 2: Example of a model of use**

As in the first step of the project we considered the open community of users in general, we have constructed generic high-level models such as "free browsing in the Website", "taxonomy-based research", "keyword research", "peer-to-peer interaction" (etc.) and mixed models such as "taxonomy-based research + peer-to-peer interaction". More task-oriented models would be models corresponding to how a specific category of people use this data, similarly to Hynecosum. Figure 2 presents the model of use "free browsing and promotion of peer-to-peer interactions". The model mixes pages proposed by the Website (states in white in figure 2; these states correspond in fact to "types of pages" such as "result from a search", i.e. not the exact Urls of the pages but a generic form of their Urls) and pages proposed by the recommender system (states in grey). We will use this example to illustrate the following subsections.

### 3.4.2. Definition of the tips and functionalities associated to the model

Once the states and links from the model of use has been identified the next step is to associate each path (i.e., each graph transition) with information and/or additional functionalities that appear useful for the considered purpose and that are not directly provided by the original Website (or not in a convenient way). The information and functionalities associated to a model are

to be defined from the information about the user's navigation that is accessible by tracking his actions in respect to the model of use (i.e., where the user is and where he was previously) and/or the requirements for this model (drawbacks of the current Website, promoters' intentions) and/or a user's profile (if any). Different models can be constructed from a set of common functionalities.
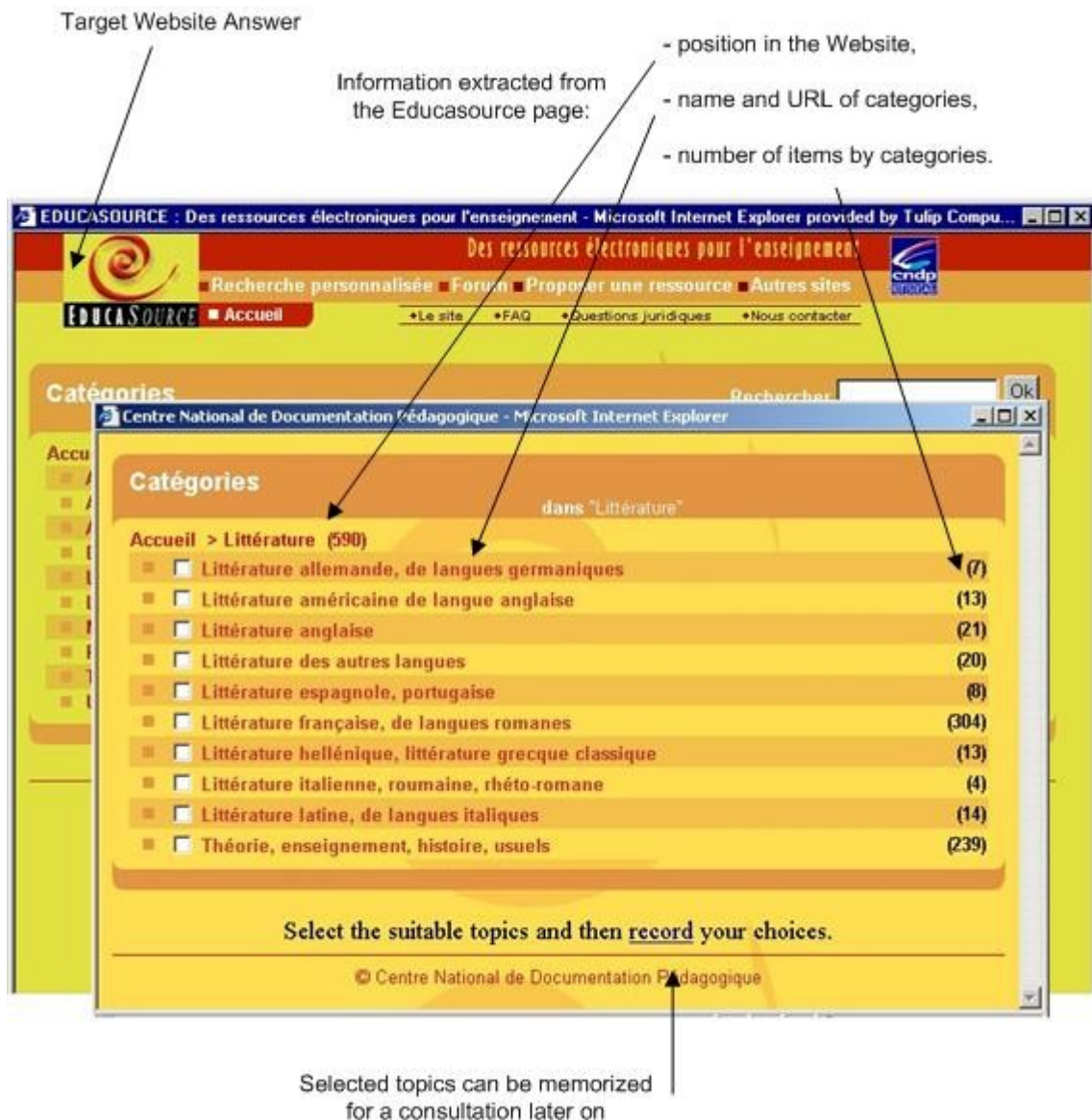
In the context of Educasource we have identified different types of functionalities such as:

- Navigation issues: highlight the different research facilities at every step; recall the different decisions that have conducted to the current step and allow moving back; highlight the current position in the navigation structure and the taxonomy hierarchy; propose memorization management facilities in order to allow focusing on search issues and analyze in detail the pages that appeared as possibly interesting later on.

- Research facilities issues: extend search where the number of results is too poor; focus search where the number of results is too big; extend search on parts of the Website that have not been visited yet; highlight links with other related Websites.

- Results management issues: structure (e.g., by domains or by keywords) the results where the number of answers is too big; allow the creation and the management of individual bags of resources (e.g., define a set of resources useful for preparing geometry exercises).

- Peer- to- peer help: allow to make one's bags of resources accessible to other users; highlight that a given resource is related to a discussion in a forum; highlight that a given resource is part of a bag created by some other user.

In the model of use presented in figure 2, the C and D states propose additional navigation functionalities: in C the user can memorize pages that seem to be interesting for a consultation later on while in D the user can access pages previously memorized. B proposes a peer- to- peer support functionality by promoting the creation and the sharing of bags of resources, a behaviour that encourages the communication between teachers.

### 3.4.3. Example

Figures 3 & 4 present two examples of the display of a window (generated by the recommender system) proposing additional functionalities on top of the original Website. Both of them are based on the model of use presented in figure 2.
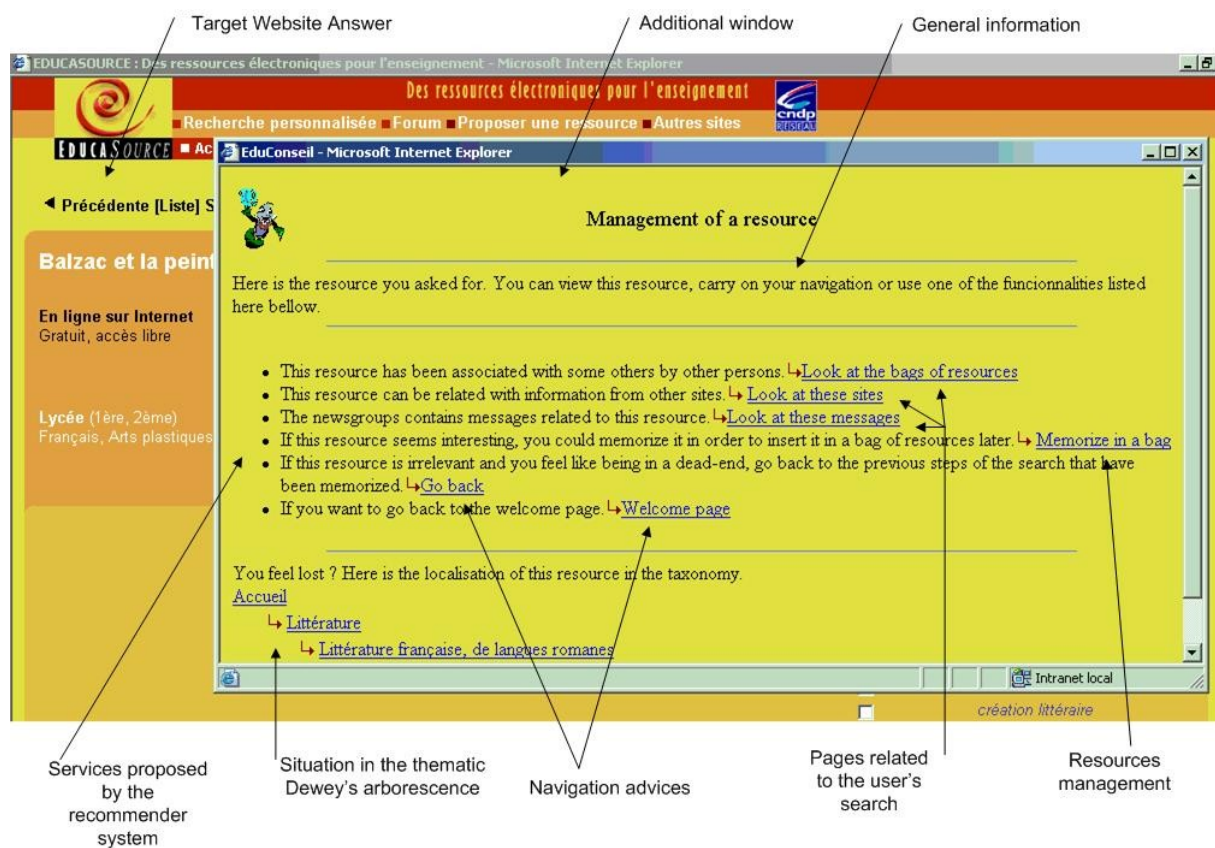
**Figure 3: Additional window generated by the recommender system in order to allow the user to select different subtopics**

In figure 3 the snapshot is taken when the user browses Dewey's taxonomy and is confronted with a list of sub-domains. This corresponds in figure 2 to an access to state "0" (action "2"). An identified difficulty of the taxonomy search is to keep in mind the overall navigation and (for example) the fact that at some point several sub-domains appeared interesting but one had to follow a unique link. With our recommender system the Website is "adapted": the user is presented (within a first window not shown in figure 3) with different functionalities including "multiple-selection", and if he selects "multiple-selection" the system displays a second window (presented in figure 3 on top of the original Website answer; this window corresponds to state "C" of figure 2) that presents the user with the possibility of memorizing the different interesting sub-domains (and he will be invited to come back to this list later on (cf. figure 4))._

In figure 4 the snapshot is taken when the user accesses to a resource descriptor. This corresponds in figure 2 to an access to states "11", "3" or "6". Here again the overall service proposed by the Website is adapted to the situation (user navigation) in respect to the model of use: the user is asked (within the window presented in figure 4) if he would like to access to some additional functionalities:
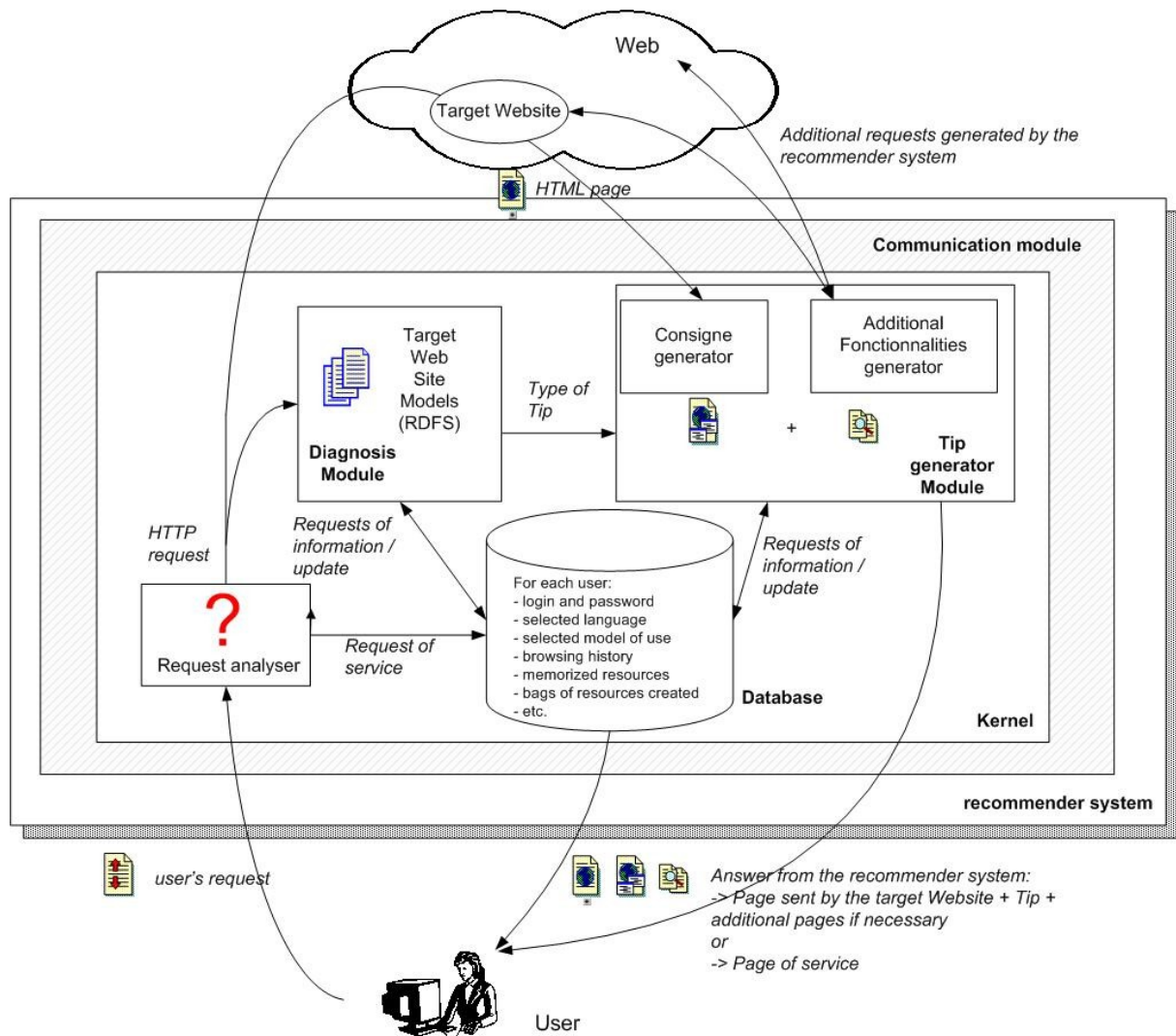
- Functionalities dedicated to the management of bags of resources (state B in figure 2).

- Functionalities corresponding to peer-to-peer issues (state D in figure 2) such as relevant pages detected in other users' bags of resources or interesting newsgroup messages (when the systems detects that the resource descriptor that has been selected by the user can be connected with some other data).

- Functionalities corresponding to navigation facilities such as allowing coming back to previously memorized interesting sub-domains (state C; cf. figure 3) or highlighting links towards referenced external Websites pages (part of state D).



**Figure 4: Additional window generated by the recommender system in order to support the management of resources**

### 3.5. Implementation of the additional functionalities

This last phase is purely technical. In some cases, customizing the Website for a given model of use only requires proposing pertinent information and tips and/or reorganizing functionalities that are available from the original Website. In some other cases, additional functionalities have to be implemented from scratch and/or by combining some of the original Website. As examples, managing bags of resources or peer-to-peer help had to be implemented from scratch, while enlarging a research is implemented by combining the original Website functionalities: the recommender system generates additional requests (in some cases, recursively), sends them to the Website, compiles the responses and proposes them to the user.



**Figure 5: Overall architecture**

Although this is the simplest case, proposing complex services obtained by combining basic functionalities from the original Website highlights the basic idea of our work. Such functionalities could be proposed directly by the Website, but functionalities that "can be of interest in certain cases" (according to the navigation and/or the Website answers) are very numerous and proposing all of them in the basic interface is not possible and not pertinent for all users. Users could combine the original functionalities themselves but

this requires skills that cannot be expected from basic users. Our approach allows adapting the original Website services runtime, according to the navigation and the considered model of use. Each model (and its associated functionalities) is studied according to a specific way of using the Website and therefore does not have to mix different (in certain cases contradictory) objectives as the original Website has (which generally conducts to a "neutral" interface that in fact denotes the underlying database). For the designers, this model- based approach is very interesting because of its flexibility: it is easy and costless to create different models or to modify them according to new emerging uses or new expectations.
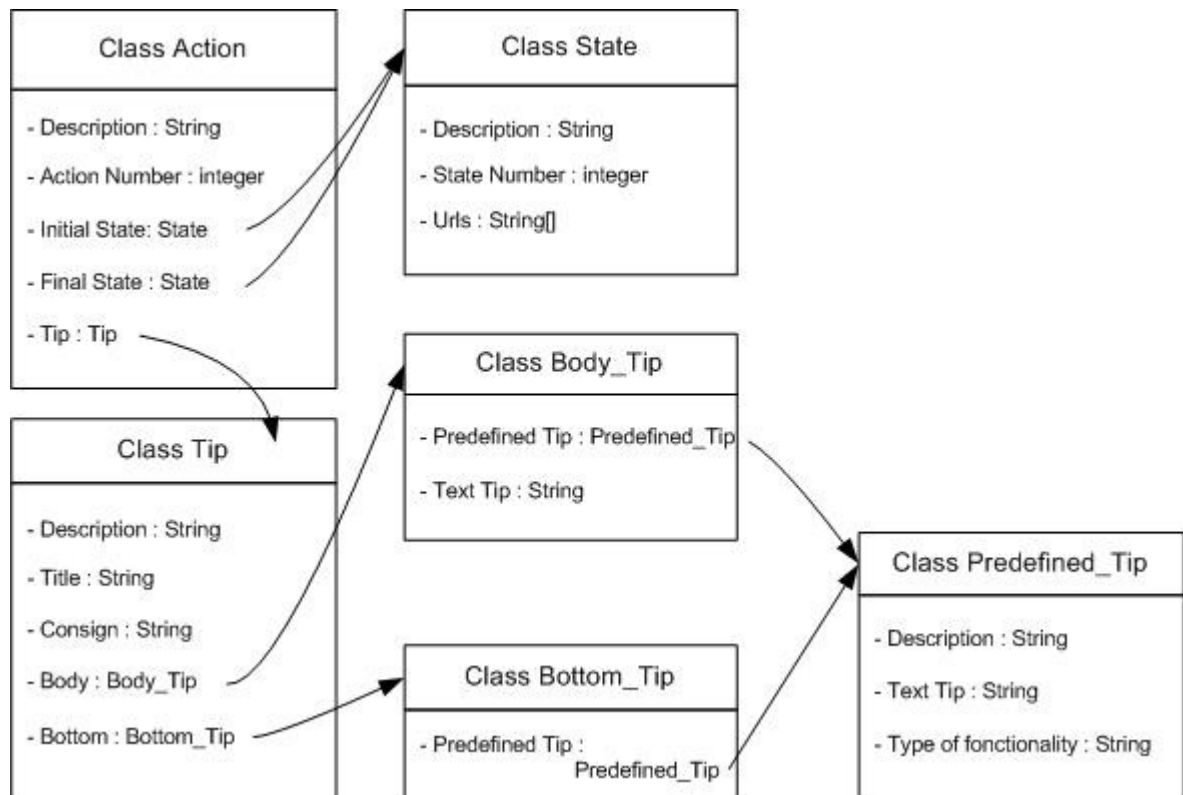
## 4. ARCHITECTURAL AND IMPLEMENTATION ISSUES

As explained previously, a key aspect of our approach is the non- modification of the original Website (epiphyte approach). This is achieved by the implementation of the overall system as a proxy- like architecture that comes in- between the user and the original Website (cf. figure 5). Within this architecture, the user is not connected to the original Website any more: he is connected to the recommender system that intercepts the data flow between the user and the original Website (HTML requests, Web pages). However, this is totally transparent for the user (and can be achieved automatically by a redirection).

The general algorithm is as follows (cf. figure 5). The user's action (i.e., the request to the server) is intercepted by the system. If a Website page is requested (in this article we are not considering service requests as connexions to the recommender system or language selections) the system forwards this action to the original Website and receives its response. Meanwhile, the recommender system diagnosis- module analyses the user's request in respect to the current model of use and deduces the corresponding action (i.e., transition of the state graph). From this action the tip generator module identifies what type of information or functionalities may be useful for the user (in respect to the current model of use). It then generates the corresponding page, using information from the original Website response (if pertinent) and (eventually) additional information (for instance, the tip generator may send itself some other requests to the original Website or to some other related Websites). The original Website response and the additional page generated by the recommender system (if any) are then sent to the user navigator. The only modification of the original Website response by the recommender system is the addition of a hidden Javascript code that will automatically open the additional window (if any). As explained previously, the recommender system acts in a separate window and does not modify the original Website pages.

Building a recommender system by instantiating this generic architecture requires the construction of the different considered models of use and their associated tips and functionalities. In order to ease this process and allow reuse we use the Protégé 2000 tool (21). This tool allows us to define generic classes to model states, actions, tips and predefined functionalities and their relations. Figure 6 presents the structure of the generic classes defined with

Protégé: the "state" class (models the Website pages, i.e., the Urls), the "action" class (models the links between the states, a link being defined by an initial and a final state) and the "tip" class (models the tips and the functionalities associated to the actions).



**Figure 6: The generic classes**

Once the generic classes are instantiated with a given model of use, the corresponding description is recorded in RDFS format (22) files using Protégé. This RDFS data is then transmitted to the recommender system and interpreted by the diagnosis module. We have chosen the RDFS format since, as a standard format, it allows compatibility and evolutivity.

Figure 7 presents an example of instantiation of the classes via the Protégé interface.
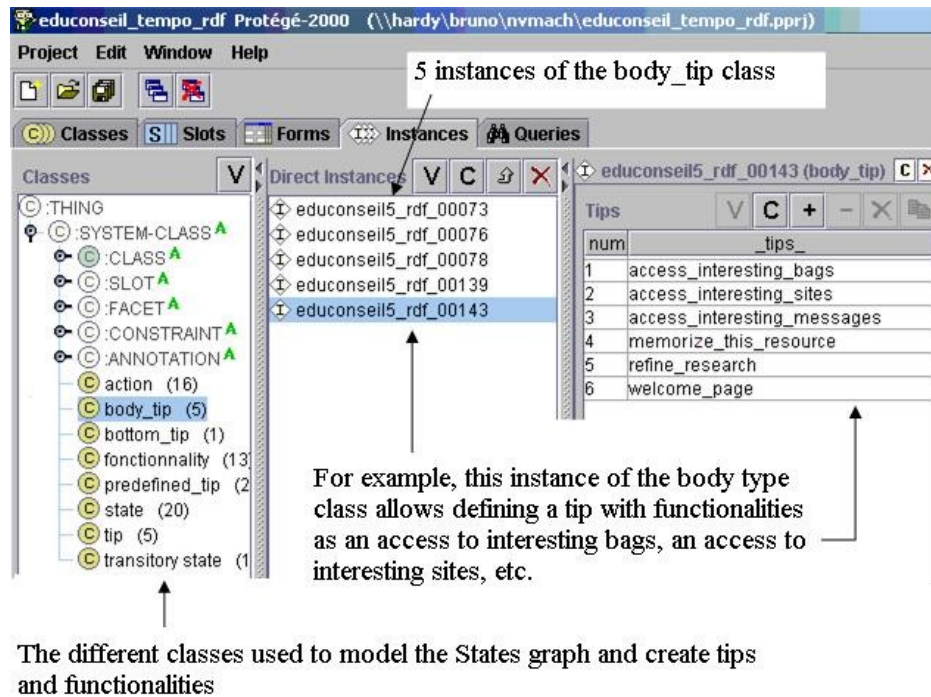
## 5. DISCUSSION

### 5.1. Evaluation

The approach we have described has been tested in the context of the development of a recommender system prototype for the French Ministry of Education and Research, experience from which we have taken the examples illustrating this paper. Although the prototype was finally unfortunately not

used (for reasons that have no relations with this project) the construction of the prototype allowed us to validate the putting into practice of the approach.

From a technical point of view, the prototype was implemented on a middle range Unix server distant from the original Website. In terms of time of response, accessing to the Website through the recommender system rather than directly did not cause any problematic delay.



**Figure 7: Protégé 2000 models**

For the user, the overall result is homogeneous and unsurprising: the original Website remains unchanged, the recommendations only appear in separate windows that can be closed at any moment, and the proposed help is predictable as it is based on an explicit model. The fact that the help appears pertinent (i.e., the "quality" of the underlying model) is an aspect that requires long-cycle experimentations we were not able to conduct yet. Here again, the fact that the underlying model of use can be easily refined leave us definitely confident.

For the modeller, constructing the models of use is a problem in itself. The difficulty depends on the complexity of the Website, its different types of users and the underlying tasks that one wants to support and/or the underlying behaviours that one wants to promote, and the explicitness of information that is available (in our case, the fact that an external audit provided part of this information was a great advantage). Although the drawbacks of Website are generally quite easy to fix, proposing a solutions that allows overcoming them is not trivial. We believe that this is however linked to the fact it is difficult to stop thinking in terms of a "unique static service" and to skip to a view in terms of "basic services + different logical

extensions". We also noticed that building such a recommender system can help simplifying the original Website as some "non basic" features can be moved to some of the recommender system models of use.

## 5.2. Scope

The approach we propose is based on the identification of prototypical uses. The scope of this model- based approach is therefore limited to domains and contexts where such "models of use" can be identified (repository- like Websites are typically good candidates). Within such contexts, a central advantage of our approach is its lightness: no modification of the original Website, no construction of a domain- ontology or metadata annotation of the Website contents (the analysis is structure- based and not domain- based), no user modelling techniques. This approach can however be enhanced by using other adaptive techniques based on domain ontologies, user modelling techniques (20), contend- based techniques (5), (23) or collaborative filtering techniques (5), (24) to propose adaptive feature to users or evolved predefined functionalities to models of use designers.

In comparison to completely automated approaches, our approach has an intrinsic cost: the construction of the models of use. However, it can be noticed that this cost is to be paid by the recommender system designer, and not by the user. This is an important difference with the approaches of the "cold start problem" where users have to participate in the knowledge acquisition phase.

## 5.3. Identification of the model of use to be used

Our approach is based on the notion of "models of use", that serves as basis for tracking the user's actions and proposing tips and additional functionalities. The identification of the model of use that should be used for a given user is therefore a central issue. Two approaches can be considered. The first one is to attempt to automatically detect the model that best denotes the user's needs on the basis of user modelling techniques, eventually skipping from one model to another when necessary. This is the usual approach to this kind of issue (25). We have decided to use an alternative approach that consists in presenting the user with different models at the beginning of the session and letting him choose the one that corresponds to his needs (eventually, none of them). Although this requires an initial interaction with the user, we believe (similarly to (18)) that this approach is (when adequate) much simpler and much more powerful. User modelling techniques are often imprecise and it is well known that submerging users with unsuitable information and functionalities is disastrous. Asking the user to select a model creates a completely different context as it results in an explicit contract between the system and the user: "I declare that my objective is just to browse through the Website and have a look at what my peers associate to the resources that interest me and I agree that the system is going to support me for this task". At any moment it is possible to skip to another model or to deactivate the recommender system.

### 5.4. An epiphyte approach

From an architectural point of view, our "epiphyte approach" presents several advantages. First, and this is a central issue, it is possible to add a recommender system to an existing Website without modifying it nor accessing to its logs or its server. Second, by focusing on "models of use" and not on the target Website one avoids difficulties and biases that occur when attempting to make a meta- level system (here, the diagnosis module) that attempts to access model- level issues of an object- system (here, the target Website structure and underlying expectations) whilst only having access to its implementation (this type of issue has been extensively studied in (26)). Third, adaptivity is addressed by constructing models that act as "logical extensions" of the target Website. This allows focusing on this issue *per se* and with an extended flexibility (easy definition, test, repair of the models; easy evolution if the Website evolves, new types of users are targetered or new uses emerge, etc.). Fourth, the overall architecture is generic and, from software and knowledge engineering points of view, easy to maintain, extend and improve.

## 6. CONCLUSIONS

We have presented an approach that allows enhancing the adaptivity of an existing Website by the means of a recommender system. The recommender system is designed as an epiphyte system that is plugged onto the target Website and intercepts the communications between the user and the target Website, which is not modified. It adapts the original Website response by displaying additional information, tips and functionalities in a separate window. These features are dynamically generated from the analysis of the user's navigation in respect to predefined prototypical uses of the Website. This approach can be used as an alternative and/or in association to other approaches related in the literature.

The putting into practice of our approach requires two main efforts: (1) building of the models of use (from an analysis of the Website and its effective use) and (2) implementation of the (eventual) additional functionalities provided by the system. Different works attempt to support users whilst avoiding such issues, for instance with content based (5) or collaborative filtering techniques (7). We however believe that these efforts are the worthwhile price to pay to enhance a Website adaptivity by taking into account the user's (and/or Website promoter's) intentions. We are currently working on the design of tools that will ease these efforts. For instance, our architecture will benefit from its association with tools that can explore a target Website and build the underlying graph such as Webbot (27), WWWPal (28) or Web2Rfs (29). Another direction of work is to consider Website patterns. It has been shown in (30) that "developers use recurrent patterns when designing Websites" (corresponding to entire Websites or parts of Websites), patterns that reflect prototypical types of navigation (browsing a site in depth: "tree pattern"; browsing the pages returned by a search engine: "indexed sequence

pattern"; etc.). It is therefore possible to consider generic recommender systems associated to generic patterns, which would limit the efforts when building a given instantiation. From another point of view, it is also possible to consider the recommender system issues while designing the original Website. Although an advantage of our approach is to allow enhancing the adaptivity of existing Websites, it can also be used, in parallel, to build a basic static Website and different "logical extensions" proposed by the means of a recommender system.

## 7. REFERENCES

(1) SCHRAEFEL, M.C., KARAM, M. and ZHAO, S., "mSpace: interaction design for user- determined, adaptable domain exploration in hypermedia". In *Proceedings of the AH2003 Workshop*, 2003.

(2) AROYO, L., DE BRA, P. and HOUBEN, G.J., "Embedding information retrieval in adaptive hypermedia: IR meets AHA!". In *Proceedings of the AH2003 Workshop*, TU/e CSN 03/04, pp. 63- 76, Budapest, Hungary, May 2003.

(3) CHEN, C.C., CHEN, M.C. and SUN, Y.S., "PVA: a self- adaptive personal view agent system" In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 257- 262, 2001.

(4) BAILEY, C., EL- BELTAGY, S.R. and HALL, W., "Link augmentation: a context- based approach to support adaptive hypermedia". In *Proceedings of the 3rd Workshop on Adaptive Hypertext and Hypermedia*, August 14- 18, Århus, Denmark, pp. 55- 62, 2001.

(5) MOONEY, R. and ROY, L., "Content- Based book recommending". In *Proceedings of the SIGIR- 1999 Workshop on Recommender Systems: Algorithms and Evaluation*, Berkeley, 1999.

(6) GARLATTI, S. and IKSAL, S., "A semantic Web approach for adaptive hypermedia". In *AH 2003: Workshop on Adaptive Hypermedia and Adaptive Web- Based Systems*, Budapest, Hungary, Johnstown Pennsylvania, USA and Nottingham, UK, 2003.

(7) LIN, W.- Y., ALVAREZ, S.A. and RUIZ, C., "Efficient adaptive- support association rule mining for recommender systems". In *Data Mining and Knowledge Discovery*, Vol. 6, No. 1, pp. 83- 105, Jan. 2002.

(8) BRUSILOVSKY, P., "Adaptive educational hypermedia". In *Proceedings of PEG'2001*, 8- 9. Tampere, Finland, 2001.

(9) AH2003, Proceedings of the AH2003 Workshop, Budapest, Hungary, May 2003, Johnstown, USA, June 2003, Nottingham, UK, August 2003.

(10) FRASINCAR, F., HOUBEN, G.J. and VDOVJAK, R., "Specification framework for engineering adaptive web applications". In *WWW2002, The Eleventh International World Wide Web Conference, Web Engineering Track*, Honolulu, Hawaii, USA, 7- 11 May 2002.

(11) CARR, L., BECHHOFER, S., GOBLE, C. and HALL W., "Conceptual linking: ontology based open hypermedia". In *Proceedings of the 10th*

*International World Wide Web Conference*, Hong Kong, China, pp 334-342. ACM Press, 2001.

(12) MIDDLETON, S.E., ALANI, H., DE ROURE, D. and SHADBOLT, N.R., "Exploiting synergy between ontologies and recommender systems". In *The Eleventh International WWW Conference, Semantic Web Workshop 2002*, May 2002, Hawaii, USA.

(13) MIDDLETON, S.E., DE ROURE, D. and SHADBOLT, N.R., "Capturing knowledge of user preference: ontologies in recommender systems". In *Proceedings of the First International Conference on Knowledge Capture (K-CAP 2001)*, Victoria, B.C. Canada, 2001.

(14) ARJONA, J.L., CORCHUELO, R. and Toro, M., "Automatic extraction of information from the Web". In *proceedings of IDEAS'02. 5th Iberoamerican Workshop Requirements Engineering and Software Environments* pp. 25-35. 2002. ISBN/D.L.: 959-7160-14-5, La Havane, Cuba, 2002.

(15) PAQUETTE, G. and TCHOUNIKINE, P., "Towards a knowledge engineering method for the construction of advisor systems". In *Proceedings of the International Conference on Artificial Intelligence in Education*, 1999, Le Mans (France), pp. 753-755.

(16) PAQUETTE, G., PACHET, F., GIROUX, S. and GIRARD, J., "EpiTalk: generating advisor agents for existing information systems". *Journal of Artificial Intelligence in Education 7* (3/4):349-379, 1996.

(17) RICHARD, B., TCHOUNIKINE, P. and JACOBONI, P., "An architecture to support navigation and propose tips within a dedicated Website". In *Proceedings of the Web Intelligence Conference*, IEEE, p 278-284, Halifax, Canada, 2003.

(18) VASSILEVA, J., "A task-centered approach for user modeling in a hypermedia office documentation system". User Model. User-Adapt. Interact. 6(2-3): 185-223 (1996).

(19) CHAMPIN, P.A., PRIÉ, Y. and MILLE, A., "MUSETTE: Modelling USEs and Tasks for Tracing Experience", Proc. *From structured cases to unstructured problem solving episodes* - WS 5 of ICCBR'03, Trondheim (NO), NTNU, Trondheim (NO), pp279-286, June 2003

(20) BRUSILOVSKY, P., Adaptive hypermedia. In *User Modelling and User Adapted Interaction*, 11, 1-2, pp. 87-110, 2001.

(21) Protégé 2000. http://protege.stanford.edu .

(22) BRICKLEY, D. and GUHA, R.V, "RDFS: Resource Description Framework Schema". In *W3C Working Draft* 12 November 2002, http://www.w3.org/TR/2002/WD-rdf-schema-20021112/

(23) DESMONTILS, E. and JACQUIN, C., "Indexing a web site with a terminology oriented ontology". In *The Emerging Semantic Web*, I.F. Cruz, S. Decker, J. Euzenat and D. L. Mc Guinness Ed., IOS Press, pp. 181-197, 2002. (ISBN 1-58603-255-0).

(24) BUONO, P., COSTABILE, M. F., GUIDA, S. and PICCINNO, A., "Integrating user data and collaborative filtering in a Web recommendation system". In *Hypermedia: Openness, Structural Awareness, and Adaptivity*, Reich, S., Tzagarakis, M.M., and Debra, P.M.E. (Eds.), Lecture Notes on Computer Science, Vol. 2266, Springer, February 2002, 315- 321.

(25) WI 2003, *Proceedings of the Web Intelligence Conference*, IEEE, Halifax, Canada, 2003.

(26) REINDERS, M., VINKHUYZEN, E., VOSS, A., AKKERMANS, H., BALDER, J., Bartsch- SPÖRL, B., BREDEWEG, B., DROUVEN, U., VAN HARMELEN, F., KARBACH, W., KARSEN, Z., SCHREIBER, G. and WIELINGA, B. (1991). "A conceptual modelling framework for knowledge- level Reflection". *AI Communications* 4 (2/3):74- 87.

(27) http://www.w3.org/Robot/

(28) PUNIN, J. and KRISHNAMOORTHY, M., "WWWPal system - A system for analysis and synthesis of Web pages". In *Proceedings of the WebNet 98 Conference*, Orlando, November, 1998.

(29) MARTIN, J. and MARTIN, L., "Web site maintenance with software- engineering tools". In *Proceedings of the 3rd International Workshop on Web Site Evolution (WSE'01)*, Florence, Italy, 2001.

(30) RICCA, F. and TONELLA, P., "Understanding and restructuring Web sites with ReWeb", *IEEE MultiMedia*, vol.8, n. 2, pp. 40- 51, April- June 2001.