Combining Social- and Information-based Approaches for Personalised Recommendation on Sequencing Learning Activities

Citation for published version (APA):

Hummel, H., Drachsler, H., Janssen, J., Nadolski, R., Koper, R., Berlanga Flores, A. J., & Van den Berg, B. (2007). Combining Social- and Information-based Approaches for Personalised Recommendation on Sequencing Learning Activities. *International Journal of Learning Technology*, *3*(2), 152-168. https://doi.org/10.1504/IJLT.2007.014842

DOI:
10.1504/IJLT.2007.014842

Document status and date:
Published: 01/01/2007

Document Version:
Peer reviewed version

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

https://www.ou.nl/taverne-agreement

Take down policy
If you believe that this document breaches copyright please contact us at:

pure-support@ou.nl

providing details and we will investigate your claim.

Downloaded from https://research.ou.nl/ on date: 04 Nov. 2024

Open Universiteit
www.ou.nl

Combining social-based and information-based approaches for personalised recommendation on sequencing learning activities

Hans G.K. Hummel*, Bert van den Berg,
Adriana J. Berlanga, Hendrik Drachsler,
José Janssen, Rob Nadolski and Rob Koper

Educational Technology Expertise Centre (ETEC)
Open University of the Netherlands (OUNL)
Valkenburgerweg 177
6419 AT Heerlen, The Netherlands
Fax: +31 45 5762 907
E-mail: hans.hummel@ou.nl
E-mail: bert.vandenberg@ou.nl
E-mail: adriana.berlanga@ou.nl
E-mail: hendrik.drachsler@ou.nl
E-mail: jose.janssen@ou.nl
E-mail: rob.nadolski@ou.nl
E-mail: rob.koper@ou.nl
*Corresponding author

Abstract: Lifelong learners who select learning activities to attain certain learning goals need to know which are suitable and in which sequence they should be performed. Learners need support in this way-finding process, and we argue that this could be provided by using Personalised Recommender Systems (PRSs). To enable personalisation, collaborative filtering could use information about learners and learning activities, since their alignment contributes to learning efficiency. A model for way-finding presents personalised recommendations in relation to information about learning goals, learning activities and learners. A PRS has been developed according to this model, and recommends to learners the best next learning activities. Both model and system combine social-based (*i.e.*, completion data from other learners) and information-based (*i.e.*, metadata from learner profiles and learning activities) approaches to recommend the best next learning activity to be completed.

Keywords: Personalised Recommender Systems; PRS; collaborative filtering; sequencing; learner profile; learning technology specifications; domain model for way-finding; learning technology.

Reference to this paper should be made as follows: Hummel, H.G.K., van den Berg, B., Berlanga, A.J., Drachsler, H., Janssen, J., Nadolski, R. and Koper, R. (2007) 'Combining social-based and information-based approaches for personalised recommendation on sequencing learning activities', *Int. J. Learning Technology*, Vol. 3, No. 2, pp.152–168.

Copyright © 2007 Inderscience Enterprises Ltd.

Biographical notes: All authors work at the Educational Technology Expertise Centre of the Open University of the Netherlands, and are currently involved in projects researching critical facilities for way-finding in learning networks. Hans G.K. Hummel works as an Associate Professor and his main interests are focused on way-finding facilities, learning technology specification and competence-based education. Bert van den Berg works as an Educational Technologist and is involved in the research and development of tools for providing navigational support. Adriana J. Berlanga works as an Assistant Professor with a background in adaptive hypermedia research and is involved in the research and development of models for competence development. Hendrik Drachsler is a PhD student carrying out research into personalised recommendation systems. José Janssen works as an Assistant Professor and is currently carrying out PhD research into a learning path specification. Rob Nadolski works as an Assistant Professor with experience in researching competence-based multimedia and is currently focusing on competence modelling. Rob Koper works as a Full Professor and Director of the technology development programme, focusing on self-organised distributed learning networks for lifelong learning.

1 Introduction

Most curricula have been designed carefully by professionals in the field. It has been the learners' main task to follow the sequence that was designed in the curriculum. It is questionable whether available curricula contain the most suitable order for each learner. Besides, from a lifelong learning perspective, learners will not just follow available curricula. At various stages of their lives they will face the task of selecting and mixing both formal and informal learning activities, taken from different sources (like the internet, peer discussions or training courses). In the absence of any 'designed order' of a curriculum, it will be hard to select and sequence the right learning activities. Learners' problems in 'way-finding' (the process of selecting and sequencing learning activities, which we consider synonymous to 'navigation') will decrease the efficiency of education provision (the ratio of output to input) and increase the cost. For instance, Dutch Open University students reported a lack of adequate information on study possibilities at an early stage of study, and problems in getting a good overview of the number and best sequence to study modules (Joosten and Poelmans, 1998).

Even within the context of one institution, this way-finding problem seems to be caused by inadequate and incomplete information rather than by lack of information. Impersonal and inadequate course selection and sequencing guidance by institutions may be held partially accountable for early drop-out rates, as, for instance, 21% of British Open University students reported (Simpson, 2004). This way-finding problem will become even more urgent within the context of distributed lifelong learning networks (Koper and Sloep, 2003). Such networks will require learners to make well-informed choices from the vast amount of learning activities on offer from various sources. When lifelong learners can choose learning activities from a greater variety from different providers, traditional institutional facilities like course catalogues or face-to-face study advice no longer offer adequate guidance.

Learners need personalised advice when (suitable) curricula are not available. Although research reveals a relation between advice and drop-out rates, advice appears to be just one of many factors (Rovai, 2003). There may be other alternatives to costly face-to-face advice (*e.g.*, domain-specific diagnostic self-tests), but this article focuses on Personalised Recommender Systems (PRS) as a promising solution. We explore the potential of such systems, which are based on collaborative filtering information from other learners (also called indirect social navigation) in combination with information about learning activities and learners (*e.g.*, needs and preferences). These systems provide learners with individualised way-finding advice on suitable learning activities and paths towards certain learning goals, like the attainment of competences. This social-based approach and the combination with the information-based approach have hardly been applied in learning (Herlocker *et al.*, 2004).

The second section of this article describes related work on recommender systems, and discusses which learning technologies and information-matching techniques are needed to enable personalisation. The third section presents a model for way-finding in learning, with personalised recommendations based on social- and information-based approaches, as well as a PRS that was developed according to this combined approach. The feasibility and future research issues when combining these prediction techniques will be discussed in the fourth and final section.

2 Personalised recommendation on sequencing learning activities

We start this section by describing related work on personalised recommendation (Section 2.1). Current limitations will be addressed, and personalisation will be proposed as a (partial) solution (Section 2.2). Learning technologies (Section 2.3) and information-matching techniques (Section 2.4) that would enable such personalisation are then discussed.

2.1 Related work

Most readers will already be familiar with recommender systems that offer advice to potential online buyers of books, movies or music (*e.g.*, amazon.com). Such applications are based on the collaborative filtering of information obtained from other buyers ('others that bought this book, also bought these books:'). Some even take ratings or tagged interests by individual users into account ('others that like Tarantino as director, also like these directors:'). Review studies do not mention personalised recommender systems in learning (*e.g.*, Herlocker *et al.*, 2004). The educational field imposes some specific demands on the advice required. The main differences between selecting books for reading and selecting learning activities for study are the degree of voluntariness (as most learning activities are required to obtain some learning goal), and the possibility to establish an explicit completion (as most learning activities are to be assessed for successful completion). Such differences impact the learner's motivation, and the way personalised recommendations for learning activities should be provided.

An exploratory study of a recommender system, using collaborative filtering to support (virtual) learners in a learning network, has been reported by Koper (2005). He simulated rules for increasing/decreasing motivation and some other disturbance factors in learning networks, using the Netlogo tool. Learners had to complete a certain set of learning activities, and after each completion were 'set' to complete the best next learning activity, based on the successful completion of next learning activities by others. Amongst other factors, the provision of this indirect social navigation accounted for about 5% to 12% of the increase in goal attainment (completion of the set), depending on the 'matching error'. This interaction effect shows that recommendations compensate for bad matching.

Closely related to this study is an experiment reported by Janssen *et al*. (2007). The authors offered learners a similar recommendation ('Most successful learners continued with Y after having completed X'). The recommendations did not take personal characteristics of learners (or possible 'matching error') into account. This indirect social navigation tool appeared to enhance effectiveness in a learning network (completion of the set of learning activities), but it did not increase efficiency (the time it took to complete them). Results showed that 40,2% of the learners who were offered recommendations completed the learning activities, whereas this portion was only 33,4% for those learners who were not offered recommendations; this difference implies that the contribution of recommendations to goal attainment was 6,8%.

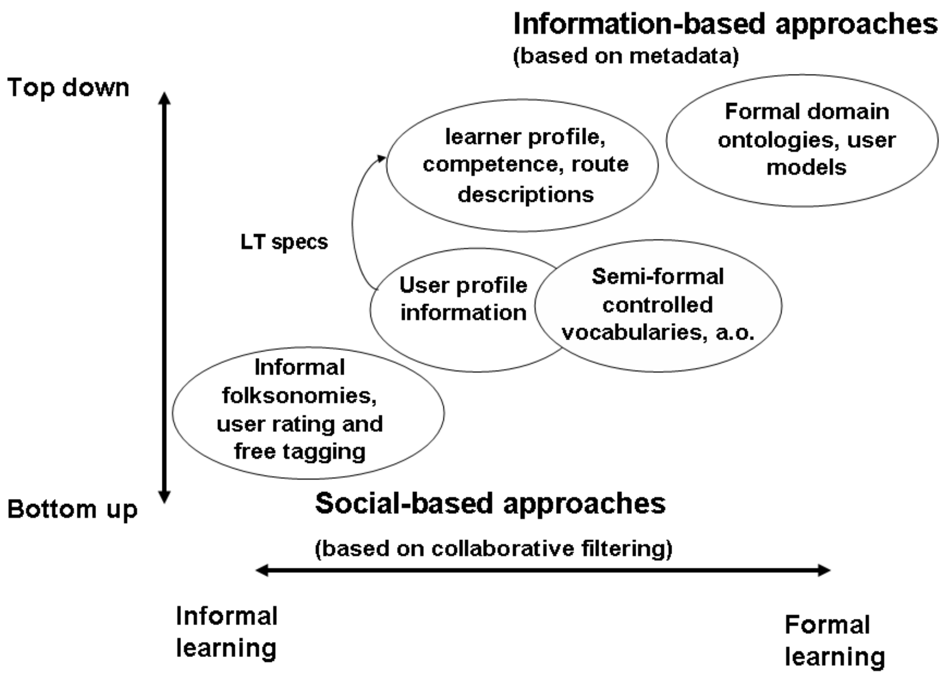
2.2 Current limitations and possible solutions

Although the first results of applying recommender systems for sequencing learning activities appear promising, the effects appear relatively small, and to base recommendations on collaborative filtering information about completions only would be too limited. In the simulation study reported by Koper (2005), 'matching error' was limited to competence levels of learners and learning activities. Mismatching can also occur on other learner characteristics such as personal needs, preferences and circumstances. For instance, Bocchi *et al*. (2004) found that student-profile characteristics accounted for about 30% of retention rates in an online MBA programme. While the aim of the Janssen *et al*. (2007) study was to recommend the most efficient learning paths (like the 'shortest route' provided by the GPS in your car), leaving the learner and learning activity characteristics out of scope, we also would need to explore which paths are most *attractive* or *suitable* (like the routes suited for bicycling). PRSs filter specific data from learning activities and learners that fit individual needs, interests, preferences or circumstances.

There are also more technical limitations to collaborative filtering. Collaborative filtering relies on large numbers of users explicitly rating or completing learning activities. When recommendations would solely depend on collaborative filtering, new or few learners (when it is hard to find similar users or when just a few users have rated the same items) would be seriously handicapped by 'cold-start' problems. Furthermore, the events registered in common log files format (as defined by W3C) are extremely low level, which complicates further analysis and more clear-cut decisions that might be required in more formal learning curricula (*e.g.*, only learning activities of certain types are allowed, or only grades above a certain threshold are considered sufficient for their completion). This makes it difficult to know which (type of) users are interacting (since only IP addresses are logged) and what (type of) interactions they are engaged in (since only URLs are logged).

One possible solution to (partially) address these limitations is to enhance the logs with additional information drawn from more or less formal descriptions about which learner did what and whether this was successful (Oberle *et al.*, 2003). Especially with a small number of learners, more information is required for exact matching of learners and activities in formal learning programmes, and more information-based (or even ontology-based) approaches will come into play. Using more structured but less formalised methods to describe learner profiles, competences and learning activities (like structured metadata or controlled vocabularies) might offer a more feasible and intermediate position. Such structured descriptions can be the basis for standardisation through open Learning Technology (LT) specifications (see Figure 1). Using LT specifications, and combining social-based with information-based approaches to match learners and learning activities, might be an approach to way-finding that addresses the cradle-to-grave challenge posed by both formal and informal lifelong learning.

Figure 1 Social-based versus information-based approaches



2.3 Learning technology specifications

Three concepts are closely related to personalised recommendation for learning in a certain domain: learner's starting position in that domain (based on learner's prior learning history and learner *profile*), the aimed *competence* profile for that domain, and the *learning path* towards that competence (also called the Competence Development Programme). Way-finding or navigation processes are about selecting and sequencing available learning activities into an individualised learning path. A learning path describes possible combinations of learning activities leading to the achievement of the competence.

Therefore, personalised recommendation in learning needs to combine and use the following pieces of the learning technology puzzle (to be further addressed in the remainder of this subsection): uniform and meaningful description of formal and informal *learning paths*; *learning activities* that are addressable and meaningfully described; uniform *learner profiles* that define needs and preferences; uniform *competence description* that defines proficiency levels that can be inferred from learner profiles; a learning path-processing engine able to compute what remains to be done by the learner to acquire the competence profile; an engine recording the completion of activities and propagating this to associated systems; and *information-matching techniques* to enable personalised recommendations. To exchange competences and recommend learning activities and learning paths to learners in an interoperable way, we will first need learning technology specifications that describe these concepts in a uniform and meaningful way.

2.3.1 Learning path description

If all providers would use a common language to describe their learning programmes and activities, PRSs could better support learners deciding between various paths to reach aimed competences. But unfortunately, no such commonly used specification exists. A number of existing approaches to (formal) curriculum modelling (*e.g.*, CDM, 2004; XRCI, 2006), together with additional requirements for a learning path description, are needed to enable personalised recommendation. Tattersall *et al.* (2007) propose IMS-LD (2003), a specification for modelling learning activities, as a strong candidate to model learning paths as well, and demonstrate that its selection and sequencing constructs appear suitable on both the level of learning activities (units-of-learning) as well as on higher levels of granularity (like competence development programmes). They note that, in addition to the curriculum-structuring concepts covered by IMS-LD, other information will be required to provide learners with more personalised advice on learning content.

2.3.2 Learner profile description

Describing and recording a learner's history and profile becomes of crucial importance to lifelong learning. Competences are not only attained through formal education, but also through work, at home or in any other context where problem solving takes place. Currently, educational providers have little possibility of adapting their formal programmes by accounting for individual needs and preferences or prior knowledge attained through other (formal) education. It becomes even harder for them if they are asked to take into consideration competences acquired through less formal or informal learning. Uniformity in assessment (to measure and accredit) and e-portfolios (to store) to enable this in the future constitute a research field on their own. The IMS Learner Information Package (IMS-LIP, 2001) and ePortfolio (IMS-ePortfolio, 2004) specifications ensure the exchange of learner records, by linking to produced artefacts and formal achievement records like references. Fields describing learner information are open and optional, so these specifications do not provide classifications of specific learner information that are easy to interpret uniformly.

2.3.3 Competence description

Cheetham and Chivers (2005) define competences as the integrated application of knowledge, skills (or competences), experience, contacts, external resources and tools to solve problems at a certain level of performance in a certain occupation or any other context. According to this definition, competence is related to three dimensions: the *type* of competence (*i.e.*, cognitive, functional, personal, professional or ethical); an occupation or performance *context*; and the proficiency *level* of a person with respect to an occupation or context. The IMS Reusable Definition of Competency or Educational Objectives (IMS-RDCEO, 2002) and the IEEE Reusable Competency Definitions (IEEE-RCD, 2006) specifications do not have much semantic value, since they simply attach IDs for a registry and URIs, and reference only to more controlled models. The HR-XML consortium (2006) goes one step further in the description of a competence and includes classes (as free text fields) where 'evidence' (with reference to learner portfolio) and 'proficiency level' (complexity, intensity, quantity) *can* be described. We suggest that a competence description should at least include an interoperable classification of *type* of competence and proficiency *level*. Sicilia (2005) proposes the formalism of ontologies to express more details in competence schemas, in order to be connected to learner profiles and learning activities. Matching learning activities to the right learners requires ontological structures and meaningful descriptions of competences in a registry (Ng *et al*., 2005).

2.4 Information-matching techniques

Information-matching or recommendation techniques work with two entities: users and items. Elements of both entities are associated with a profile carrying certain characteristics. The utility of an item (*i.e.*, learning activity) to a user (*i.e.*, learner) is usually represented by a rating function *R: Users x Items* → *Ratings*, and recommendations are estimated ratings for items which have not been 'seen' (*i.e.*, enrolled, rated, successfully completed) by the user. Within information-matching techniques a distinction is made between *information-based approaches* (based on learning technology standardisation, metadata and semantic web efforts) and *social-based approaches* (based on data mining, social software and collaborative filtering) (Balabanovic and Shoham, 1997; Van Setten, 2005). Main benefits and drawbacks of these approaches for personalised recommendation will now be described, although an in-depth treatment of specific prediction techniques and algorithms involved has to remain out of scope here. Figure 1 depicts the relation between the two approaches, the type of learning (formal versus informal) and the formalisation of information (imposed from the top downwards versus emerging from the bottom upwards).

2.4.1 Information-based approaches

Information-based approaches may use certain keywords or metadata which hold knowledge about certain characteristics of the learners and learning activities. The system then keeps track of items the user was previously interested in, and recommends items with similar or related keywords or metadata. Similarity of items is calculated with techniques based on item-to-item correlation that may use keywords in documents (Schafer *et al*., 1999). Ontological modelling has the immense (potential) advantage of exact-matching competence descriptions in the learner profile (user model) with available competence development programmes (domain model), sharing a common understanding in a machine-readable way. Exact matching is often required in more formal learning situations.

Examples of modelling learners, tutors and learner-tutor interaction can be found in Intelligent Tutoring Systems (ITS), where researchers have strived to provide better sequences of tutoring actions by including 'instructional planning' mechanisms (*e.g.*, VanLehn, 1988; Reiser and Dick, 1996). Knowledge planners use such models to input planning schemes that control decision making. Planning then becomes deciding on which schema, operators and prerequisites to use. Although such knowledge planners would use both top-down and bottom-up information (from actual user behaviour), we consider this approach to be mainly information based since there is an intimate relationship between the information used to make instructional planning decisions and the learner model (Frasson *et al.*, 1992).

A serious *drawback of modelling* is the enormous amount of work in enriching resources with metadata and the arbitrary character of such models. For example, the main problems in instructional planning were caused by limitations to *learner* modelling. A second limitation is the inability to analyse and capture all important characteristics of the *content* (*e.g.*, how to distinguish an excellent from a poor article that are using the same keywords). Most categories we employ in life are based on fuzzy concepts rather than on objective rules (*e.g.*, Where's the line between good and excellent? Do we really care about the subtle distinctions in wine-tasting?) (Morville, 2005). A third drawback is caused by the fact that words in the form of metadata are ambiguous. Our language is filled with synonyms, homonyms, acronyms and even contranyms (words with contradictory meanings in various contexts). In retrieval, the forces of discrimination and description are battling, with full-text search being biased towards description (finding general words with many meanings) (like with Google; see Brin and Page, 1998), and unique identifiers (like with an ID number of each competence in a registry; see IMS-RDCEO, 2002) offering perfect discrimination but no descriptive value whatsoever.

It will be neither possible nor necessary to fully model learning networks, but certain levels in formalising information can be distinguished. *Ontologies* describe how concepts of the world are related and represented using formal relations. An ontology is a rather strict formalisation into a machine-readable format consisting of entities, attributes, relationships and axioms (Guarino and Giaretta, 1995). *Taxonomies* can be considered a special kind of ontology. They are hierarchical structures of names and descriptions. There are also semiformalisations holding the middle ground: *structured metadata* fields (with headers like title, domain, provider of the activity), *facetted classifications*, which permit the assigning of multiple classifications to an object and the accommodating of way-finding that varies by user and task, and *controlled vocabularies* (*e.g.*, fixed categories, keyword lists, audiences) that try to control the language ambiguity.

In learning networks the information-based approach to way-finding has remained scarce. Buzza *et al.* (2004) propose a controlled vocabulary to search for learning designs from a limitless collection of units-of-learning modelled in IMS-LD. Their prototypical learning designs search engine can search on keywords, discipline, delivery mode (*i.e.*, face-to-face, online, blended) and educational rationale (*i.e.*, anchoring knowledge, developing motivation, applying theory, monitoring comprehension, adapting to task difficulty, collaborating, engaging in self-evaluation, reflecting, thinking critically). A good starting point for a facetted classification for learning could be existing frameworks (*e.g.*, offered by the AERA or the APA).

2.4.2 Social-based approaches

The big advantage of social-based approaches is that they are completely independent of the representation of knowledge in domain or user models. Instead of recommending to a specific user the items similar to his/her previously liked items, this approach recommends the items liked by other users in similar situations or with similar preferences ('peer groups'). It uses peer opinions to predict the interests of others, and matches users against the database to discover historically similar tastes. It avoids the enormous amount of work involved in enriching resources.

A serious *drawback of recommendations based on collaborative filtering* is their limited value for new or few users, that is, when it is hard to find similar users or when just a few users have rated the same items, or when no content is available about already attained competences. Clustering may then be an alternative to solve this sparsity problem (Agarwal *et al*., 2006). Collaborative filtering techniques also suffer from serious limitations when exact matching is required, and when competences in domains go beyond the verbal realm (*e.g.*, hard-to-express communication or motor skills). Although collaborative filtering has been considered a mainstream technique for recommender systems, applications with actual learning-behaviour data for recommending learners have remained scarce.

In order to better match learners and learning activities, it is possible to compute the similarity between pairs of learners and recommend other yet unaccessed learning activities. For instance, Li *et al.* (2005) combine item-based and user-based collaborative filtering, based on content information and ratings at the same time, which makes it possible to alleviate both sparsity and cold-start problems at the same time. For instance, when a database of learning activities contains both the ratings of peers (user-based) and tags describing pedagogical taste (item-based, *e.g.*, programmed instruction or problem-based learning), it becomes possible to personally recommend items targeting the taste of users, which will be apparent from their history of rating.

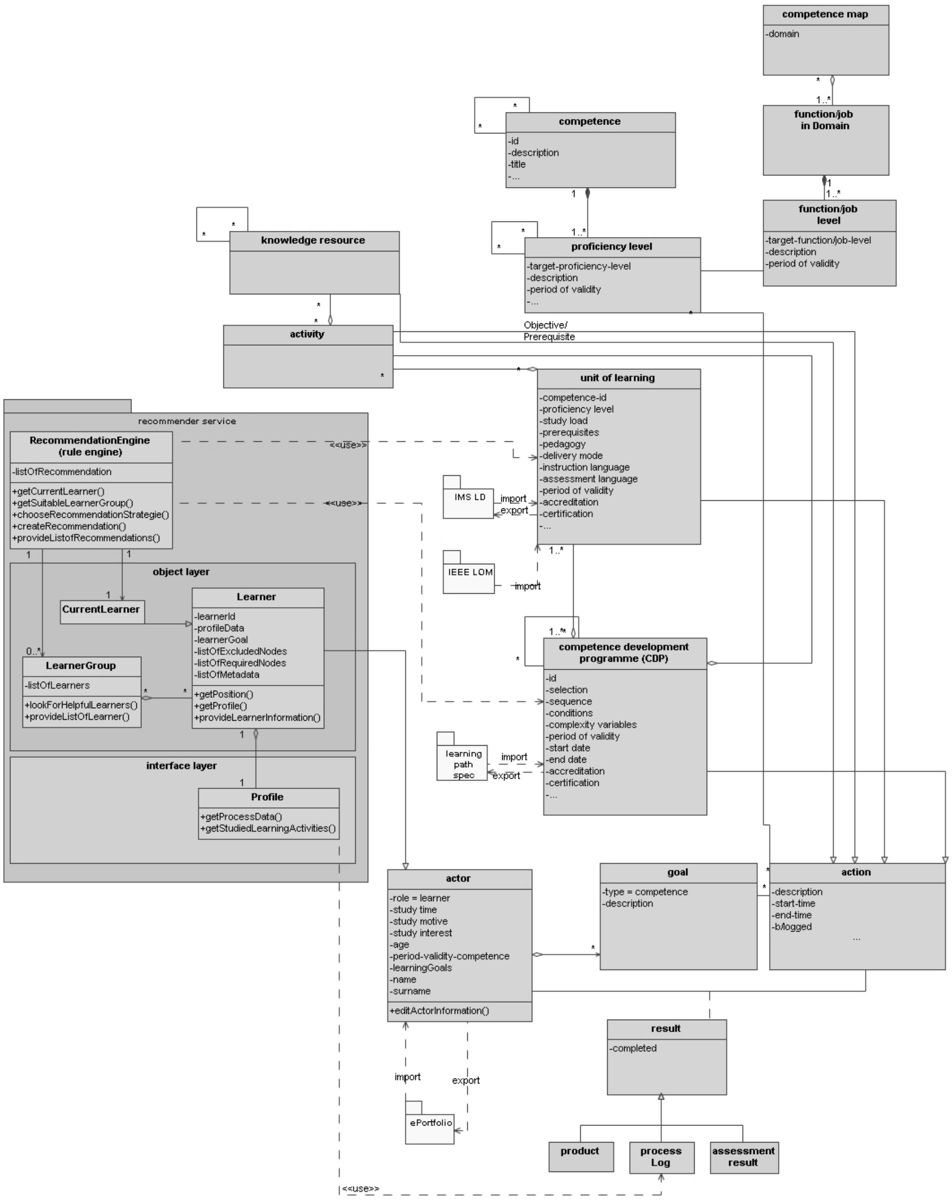
3 Model and personalised recommender system for sequencing learning activities

As described before, the goal of a PRS for way-finding in a learning network is to support the learner to *compose the most suitable sequences of learning activities* to attain competences. A model for way-finding (Section 3.1) and a PRS implemented according to this model (Section 3.2) will be presented.

3.1 A model for way-finding

A model for way-finding in learning networks is depicted in Figure 2. The model is called initial because concrete attributes of the classes are to be decided upon and validated (denoted by '…'), as well as the interfaces with services and specifications. The model is a partial elaboration of the more general TENCompetence domain model,[1] and focuses on the most relevant classes for personalised recommendations on sequencing.

Figure 2 Initial model for adaptive way-finding in learning networks



According to the model (classes between square brackets), an [Actor] (the role of which can be 'learner') has a learning [Goal] (of the type 'competence') that makes him/her perform actions. Characteristics of an actor (like study time, study motive, study interest) can be modelled according to the IMS-ePortfolio (2004) specification. [Actions] can be aimed at the acquisition of certain competences. When such an action is completed with an [Assessment Result] available, one can infer whether the [Competence] has been mastered at a certain [Proficiency Level]. Actions can include learning paths (named [CDP]), [Units of Learning] (modelled in IMS-LD), [Activities] (not modelled in IMS-LD) and [Knowledge Resources]. A unit of learning depicts the flow of activities that a learner needs to follow in order to reach certain learning goals, giving certain prerequisites. (Please note that in the remainder of this article, we simply use the term 'learning activities' to denote both units that are and are not modelled in *IMS-LD*.) A learning path is an ordered set of actions, activities and/or units of learning that can be followed in order to acquire a competence with a certain proficiency level. Learning paths can include selections or sequences of 'learning activities' (units of learning or activities), as well as conditions for their composition; attributes of a learning path can be modelled according to a *learning path specification*.

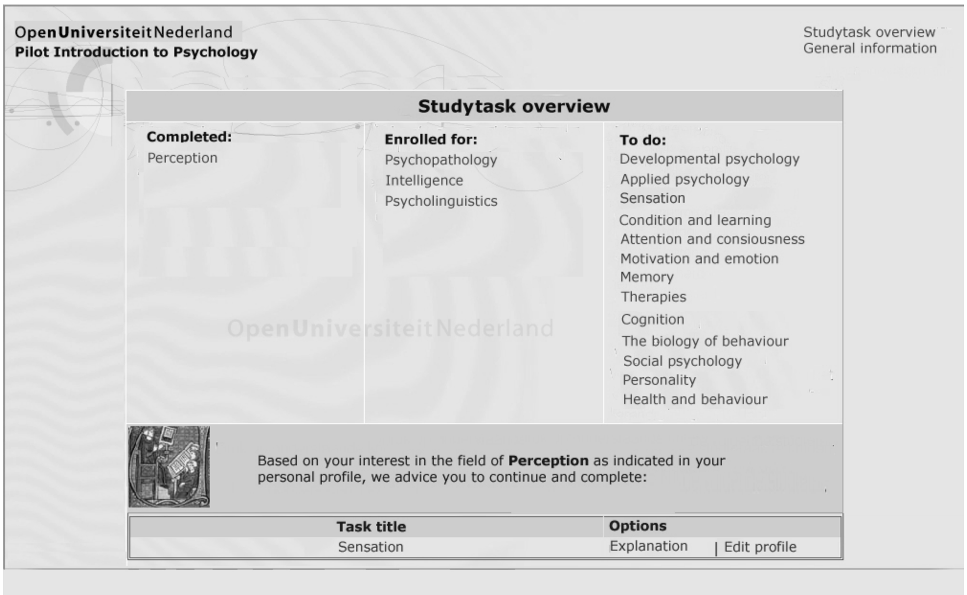
The core of this way-finding model is constituted by the [Recommender Service], where relevant information (from the classes just described) will be updated, integrated and matched to create personalised recommendations. Depending on the learning situation (*e.g.*, formal versus informal learning), a 'recommendation strategy' can be selected (*e.g.*, more or less formalised information) by a [Recommendation Engine]. This engine takes into account learner positioning and profile information for each [Current Learner] (*information-based*) and/or for each [Learner Group] (*social-based*). The [Learner Group] class calculates and generates data about similar learners, in order to provide the engine with necessary matching information about the 'peer group'. The [Learner] class gathers all required data about learners (actors with this role) from an interface layer to a *positioning service* (left out of the scope of this model, leaving the feedback from the [Position] class null) and the [Profile] information. All positioning and profile information will be continuously updated when learning activities are completed (using the [Process Log] information) or when the ePortfolio is changed. The learner data is either (subjectively) provided or adapted by the learner, or (objectively) collected by the TENCompetence system.

3.2 Experimentation with the system

A first 'proof of concept' version of a PRS has been developed according to the combined approach and model sketched above. This system currently under study does not yet apply rating or free tagging of learning activities, and is limited to fixed learner-profile metadata in a simple ontology. The system stores fixed metadata about the subdomain of the learning activities. Figure 3 shows the study task overview screen of the system with information about learning activities and an advice button for a personalised recommendation on the best next learning activity (*i.e.*, study task) to complete.

The *study task overview* shows to the students the learning activities (which have been implemented as Moodle courses) they are not yet enrolled for, those they have already started studying and those they have already successfully completed (in three columns). The current implementation of the system is limited to a fixed set of 18 formal learning activities that constitute an Introductory Psychology course from one provider. On average, each of these learning activities will take students around 13 hours to study, so students need around 240 hours to attain the aimed competence profile on the introductory level. The *personalised recommendation* for the best next learning activity to take will take *learner profile* information into account. In their profile, learners must select from available structured metadata on three learner characteristics that constitute a simple learner ontology. We have chosen to experiment with factors that have appeared predictive for study success (according to our psychology advice practice): available study time, study motive and study domain interest. On the highest level of analysis these characteristics are classified on the following three dichotomies: less or more than ten hours a week available for study; intrinsically or extrinsically motivated; and interest in 'experimental' or 'environmental' subdomains of psychology. These yield eight 'peer groups'. Collaborative filtering can be provided once results on successfully completed activities by similar peers become available.

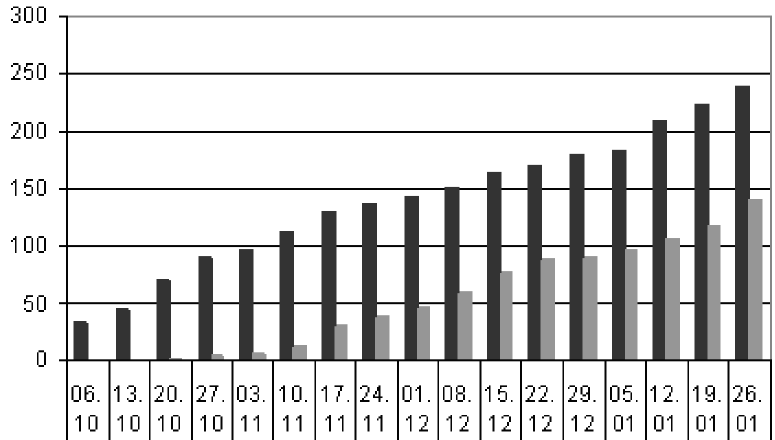
Figure 3 Personalised recommendation system: overview page



At the moment of writing (February 2007), Introductory Psychology students of the OUNL have interacted with the first pilot implementation of the system during a period of 17 weeks. An experimental group received personalised recommendations on sequencing, while an (equally sized and randomly allocated) control group received no such recommendations. At face validity, the preliminary results appear promising, although we emphasise that statistical analyses have not yet been carried out. When we consider the combination of prediction techniques, we notice a gradual increase in recommendations based on collaborative filtering, as was intended (see Figure 4). After 5, 9, 13 and 17 weeks, 7%, 25%, 33% and 37% of the recommendations were based on collaborative behaviour, with a total of 379 recommendations having been provided at the end.

When we consider the amount of completed learning activities as an important indicator of effectiveness, we notice this (cumulative) amount being consistently higher for the group receiving recommendations. After 5, 9, 13 and 17 weeks, the experimental group had completed 24%, 22%, 19% and 16% more activities when compared to the control group, with a total of 434 activities having been completed at the end of the pilot period.

Figure 4 Counts (cumulative) of recommendations based on ontology and collaborative filtering at weekly checkpoints (n = 17)



	06.10	13.10	20.10	27.10	03.11	10.11	17.11	24.11	01.12	08.12	15.12	22.12	29.12	05.01	12.01	19.01	26.01
ontology	33	45	71	90	97	113	131	137	143	151	164	170	180	184	209	224	239
collaborative filtering	0	0	2	5	7	12	30	38	47	59	77	88	90	98	107	118	140

More specific information about the method and results of this and consecutive studies will be published when experimental data have been analysed in more detail. This and consecutive versions of PRSs will be validated both in authentic field experiments and simulation studies.

4 Discussion, conclusions and future research

From a self-organisational point of view, it would be ideal if way-finding would emerge as a result of (in)direct interactions between members of the learning network, without being dependent on formalised descriptions in domain and user models. In ant colonies, a process known as stigmergy (Theraulaz and Bonabeau, 1999) is responsible for causing effective organisations and structures to emerge. Ants leave traces that influence the behaviour of others. The simulation reported by Koper (2005) and the experiment reported by Janssen *et al.* (2007) indicated that such indirect social navigation might work as the main mechanism for providing recommendations for sequencing (informal) learning activities. Limitations of these studies and collaborative filtering were discussed. The combination of social- and information-based approaches was presented in both a general model for way-finding and a specific PRS. Especially when more specific information is required for exact matching of learners and (formal) learning activities, more information-based (or even ontology-based) approaches will have to come into play. This last section discusses the feasibility of such a combination, some conclusions and future research issues.

4.1 Discussion

Research groups working on social- and information-based prediction techniques appear fragmented, and specialisation has led to a divergence of vision and vocabulary. When groups interact, they often talk past one another without translating or cooperating. The most renowned groups appear to be the group led by Tim Berners-Lee, which is working along the lines of the semantic web (top-down, provider-oriented, using ontologies and fixed taxonomies), and the group with strong advocates like Clay Shirky and David Weinberger, which works along the lines of social software and collaborative filtering (bottom-up, user-oriented, using folksonomies and free tagging).

The *feasibility* of such a complementary approach seems high when current limitations to each separate approach could be overcome. These fields of research do not necessarily have to be mutually exclusive. This complementary approach will be needed to cater to personalising recommendations in both informal and formal learning situations. Collaborative filtering will work fine for informal learning, where discrete measures and exact matching are not needed. For formal learning we need more formalised descriptions (using learning technology specification) of activities, competences and learner profiles. Metadata in the form of ontologies (semantic web) and folksonomies (social software) might bring both worlds together after all.

4.2 Conclusions

For learning networks, it was concluded that a 'hybrid and layered metadata ecology' may be ideal, where the slow layers (ontologies) provide stability and the fast layers (folksonomies) drive change. Semantic web tools and learning technology standards provide a solid semantic framework (infrastructure), where there is a need for more formal and explicit characteristics of learners and learning activities to be decided upon from the top down. Where there is a need for less formal and implicit characteristics, free tagging and folksonomies offer more flexible, adaptive user-feedback mechanisms to follow informal learning trends, emerging by serendipity and from the bottom up.

The proposed model for adaptive way-finding in learning networks seems rather general and could essentially be further instantiated according to both user-oriented and tutor-oriented paradigms, like even traditional CAI/CMI systems or ITS. For instance, Elsom-Cook (1990) described various tutorial systems along a continuum of 'guided discovery tutoring' that ranges from complete direction and control by the tutor ('Socratic dialogue' or coaching) to complete freedom of learning for the student ('learning by/while doing'). Within this continuum, our work is positioned towards the user-controlled end. Individual and collaborative information is used to generate recommendations that help users make well-informed choices. Our work does not aim to control a specific structure of collaboration (like in CSCL) or sequence of learning activities by the system (like in CAI or ITS), with a tutor/expert in charge of supervising the sequencing of activities according to the achievements of students.

According to the proposed model, we developed and tested a PRS within a formal, Introductory Psychology learning network, and preliminary results from this pilot study appear promising. We do acknowledge the limitations of focusing on a formal learning context with a fixed set of learning activities offered by one single institution. In such a situation, the learners' freedom in composing individualised selections and sequences of formal learning activities will be restricted by several (institutional) constraints and prerequisites. Besides, personalisation assumes there will be one most suitable learning path for each individual learner, whereas the personalisation process actually might be more of an optimisation process of self-guided behaviour. In more informal learning networks, the initial process of actively exploring and finding learning goals and learning opportunities might be more valuable than actually walking the learning path. Self-organisation assumes effective learning to simply emerge as active explorers learn from each other's behaviour, which includes walking dead-end streets as well. Therefore, consecutive systems are planned to broaden and generalise our findings to learning networks with contributions from various providers that entail both formal and informal learning activities.

4.3 Future research

Consecutive studies have to examine ways to retrieve learner information as effortlessly and unobtrusively as possible. We do not have to burden professional, centralised indexers when ontologies could emerge entirely locally (some parts will be shared, some not, which is no big deal), and RDF vocabularies could be freely mixed together in a pragmatic way. For instance, as opposed to the semantics-based paradigm, McCalla (2004) proposes a pragmatics-based paradigm of tagging learning activities with learner information. Each time a learner interacts with a learning activity, the (current) learner model of that learner is attached to the learning activity. Which learner data will be mined for patterns of particular use depends on the learning situation. More refined reasoning and recommendations will be possible when more and more instances accumulate. In addition to these instances, more standardised metadata could be assigned by professional indexers whenever needed. This ecological approach allows pre-assigned metadata (from ontologies like IEEE-LOM) to be refined or changed based on inferences from end use (from folksonomies tagged by learners).

Our roadmap for future research on this topic aims to also combine collaborative filtering with:

- information that can emerge bottom-up by using rating or free tagging (folksonomies) of learning activities
- information that can be distilled from more formalised descriptions (specifications) of competences, learner profiles and learning paths.

Future research will also need to establish the added value of personalising recommendations. We need to actually measure the increase in progress (effectiveness), decrease in study time (efficiency) and appreciation of users (satisfaction). Where possible, such information could be automatically added when learners interact with learning activities. When more specific information about learners and learning activities is required, metadata should be pretagged according to some ontology of attributes. In order to collect required information, learners will have to be encouraged to (automatically) update their e-Portfolio or give ratings to (attributes of) the content. We will also need to establish the extent to which learners are willing to provide the necessary information about their personal needs and preferences. When they do not volunteer to do so or when otherwise considered necessary, the potential of incentive mechanisms (Hummel *et al.*, 2005) could be considered to stimulate updating, tagging and rating.

Acknowledgement

The authors' efforts were (partly) funded by the European Commission in TENCompetence (IST-2004-02787) (http://www.tencompetence.org).

References

Agarwal, N., Haque, E., Liu, H. and Parsons, L. (2006) 'A subspace clustering framework for research group collaboration', *International Journal of Information Technology and Web Engineering*, Vol. 1, No. 1, pp.35–58.

Balabanovic, M. and Shoham, Y. (1997) 'Fab: content-based, collaborative recommendation', *Comm. ACM*, Vol. 40, No. 3, pp.66–72.

Bocchi, J., Eastman, J.K. and Swift, C.O. (2004) 'Retaining the online learner: profile of students in an online MBA program and implications for teaching them', *Journal of Education for Business*, Vol. 79, No. 4, pp.245–253.

Brin, S. and Page, L. (1998) 'The anatomy of a large-scale hypertextual web search engine', *Computer Networks and ISDN-systems*, Vol. 30, Nos. 1–7, pp.107–117.

Buzza, D.C., Bean, D., Harrigan, K. and Carey, T. (2004) 'Learning design repositories: adapting learning design specifications for shared knowledge', *Canadian Journal of Learning and Technology*, Vol. 30, No. 3, http://www.cjlt.ca/content/vol30.3/buzza.html (retrieved 3 June 2005).

CDM (2004) 'A specification of course description metadata', *Proposal 20*, October, http://cdm.utdanning.no/cdm/cdm-2.0.1/doc/courseDesc201004.pdf (retrieved 22 March 2006).

Cheetham, G. and Chivers, G. (2005) *Professions, Competence and Informal Learning*, Edward Elgar Publishing.

Elsom-Cook, M. (1990) 'Guided discovery tutoring', in M. Elsom-Cook (Ed.) *Guided Discovery Tutoring – A Framework for ICAI Research*, London: Paul Chapman Publishing, pp.3–23.

Frasson, C., Gauthier, G. and McCalla, G.I. (1992) 'Intelligent tutoring systems', *Proceedings of the Second International Conference on Intelligent Tutoring Systems, Montreal, Springer-Verlag Lecture Notes in Computer Science*, Berlin, Vol. 608.

Guarino, N. and Giaretta, P. (1995) 'Ontologies and knowledge bases: towards a terminological clarification', in N. Mars (Ed.) *Towards Very Large Knowledge Bases*, IOS Press, pp.25–32.

Herlocker, J.L., Konstan, J.A., Terveen, L. and Riedl, J.T. (2004) 'Evaluating collaborative filtering recommender systems', *ACM Transactions on Information Systems*, Vol. 22, No. 1, pp.5–53.

HR-XML (2006) 'HR-XML consortium: competency types', http://ns.hr-xml.org/2_4/HR-XML-2_4/CPO/CompetencyTypes.html (retrieved 22 March 2006).

Hummel, H.G.K., Burgos, D., Tattersall, C., Brouns, F., Kurvers, H. and Koper, R. (2005) 'Encouraging contributions in learning networks using incentive mechanisms', *Journal of Computer Assisted Learning*, Vol. 21, No. 5, pp.355–365.

IEEE-RCD (2006) 'IEEE standard for learning technology – reusable competency definition', http://ieeeltsc.org/wg20Comp (retrieved 22 March 2006).

IMS-ePortfolio (2004) 'IMS ePortfolio specification', http://www.imsglobal.org/ep/index.html (retrieved 22 March 2006).

IMS-LD (2003) 'IMS learning design specification', http://www.imsglobal.org/learningdesign/index.html (retrieved 22 March 2006).

IMS-LIP (2001) 'IMS learner information package specification', http://www.imsglobal.org/profiles/index.html (retrieved 22 March 2006).

IMS-RDCEO (2002) 'IMS reusable definition of competency or educational objective specification', http://www.imsglobal.org/competencies/index.html (retrieved 22 March 2006).

Janssen, J., Van den Berg, B., Tattersall, C., Hummel, H. and Koper, R. (2007) 'Navigational support in lifelong learning: enhancing effectiveness through indirect social navigation', *Interactive Learning Environments*, Vol. 15, No. 2, pp.127–136.

Joosten, G. and Poelmans, P. (1998) *Follow up van de eerste kennismakers: studievoortgang, plannen en doorstroom (Following up Starting Students: Progress, Plans and Flow)*, Heerlen: Open University of the Netherlands.

Koper, E.J.R. (2005) 'Increasing learner retention in a simulated learning network using indirect social interaction', *Journal of Artificial Societies and Social Simulation*, Vol. 8, No. 2.

Koper, E.J.R. and Sloep, P.B. (2003) 'Learning networks: connecting people, organizations, autonomous agents and learning resources to establish the emergence of lifelong learning', Report, Heerlen, Open University of the Netherlands, www.learningnetworks.org.

Li, Q., Myaeng, S.H. and Kim, B.M. (2005) 'A probabilistic model for music recommendation considering audio features', *Information Retrieval Technology, Lecture Notes in Computer Science (LNCS)*, Springer, Vol. 3689, pp.72–83.

McCalla, G. (2004) 'The ecological approach to the design of e-learning environments', *Journal of Interactive Media in Education*, Vol. 7.

Morville, P. (2005) *Ambient Findability*, Sebastopol, California: O'Reilly.

Ng, A., Hatala, M. and Gasevic, D. (2005) 'Ontology-based approach to learning objective formalization', http://www.sfu.ca/lkng/fina-call_for_chapter.pdf (retrieved 20 November 2005).

Oberle, D., Berendt, B., Hotho, A. and Gonzalez, J. (2003) 'Conceptual user tracking', in J.S.E.M. Ruiz and P.S. Szczepaniak (Eds.) *Web Intelligence. Proceedings of the First International Atlantic Web Intelligence Conference (AWIC 1003)*, Berlin: Springer.

Reiser, R.A. and Dick, W. (1996) *Instructional Planning: A Guide for Teachers*, Allyn and Bacon.

Rovai, A.P. (2003) 'In search of higher persistence rates in distance education online programs', *Internet and Higher Education*, Vol. 6, pp.1–16.

Schafer, J.B., Konstan, J. and Riedl, J. (1999) 'Recommender systems in e-commerce', *Proceedings of the ACM Conference on Electronic Commerce*, Denver, Colorado, pp.158–166.

Sicilia, M-A. (2005) 'Ontology-based competency management: infrastructures for the knowledge intensive learning organization', in M. Lytras and A. Naeve (Eds.) *Intelligent Learning Infrastructures for Knowledge Intensive Organisations: A Semantic Web Perspective*, IDEA Group, pp.302–324.

Simpson, O. (2004) 'Access, retention and course choice in online, open and distance learning', *3rd Eden Research Workshop*, Oldenburg, Germany, http://www.eurodl.org/materials/contrib/2004/Ormond_Simpson.html (retrieved 15 September 2006).

Tattersall, C., Janssen, J., Van den Berg, B., Hummel, H. and Koper, R. (2007) 'Using IMS learning design to model curricula', *Interactive Learning Environments*, Vol. 15, No. 2, pp.181–189.

Theraulaz, G. and Bonabeau, E. (1999) 'A brief history of stigmergy', *Artificial Life*, Vol. 5, pp.97–116.

VanLehn, K. (1988) 'Student modeling', in M.C. Polson and J.J. Richardson (Eds.) *Foundations of Intelligent Tutoring Systems*, Hillsdale, NJ: Lawrence Erlbaum Associates.

Van Setten, M. (2005) 'Supporting people in finding information: hybrid recommender systems and goal-based structuring', *Telematics Institute Fundamental Research Series No. 016*, Enschede, The Netherlands, Internal Document TI/FRS/016.

Note

1 http://dspace.ou.nl/bitstream/1820/649/13/DomainModel-version1p0.pdf