

Chapter 2

TEL as a Recommendation Context

Abstract In this chapter, we define the TEL recommendation problem and identify TEL recommendation goals. More specifically, we reflect on user tasks that are supported in TEL settings, and how they compare to typical user tasks in other recommender systems. Then, we present an analysis of existing data sets that capture contextual learner interactions with tools and resources in TEL settings. These data sets can be used for a wide variety of research purposes, including experimental comparison of the performance of recommendation algorithms for learning.

2.1 TEL Recommendation

2.1.1 Defining the TEL Recommendation Problem

In a recommender system, the items of interest and the user preferences are represented in various forms, e.g. using a single or multiple attributes for describing an item. Particularly in systems where recommendations are based on the opinion of others, it is crucial to take into consideration the multiple factors or criteria that affect the users' opinions in order to make more effective recommendations. In related research, the problem of recommendation has been identified as the way to help individuals in a community to find the information or products that are most likely to be interesting to them or to be relevant to their needs (Konstan 2004). It has been further refined to the problem (i) of predicting whether a particular user will like a particular item (prediction problem), or (ii) of identifying a set of N items that will be of interest to a certain user (top- N recommendation problem) (Adomavicius and Tuzhilin 2005). Therefore, the general recommendation problem can be formulated as follows (Deshpande and Karypis 2004): let C be the set of all users and S the set of all possible items that can be recommended. We define as $U^c(S)$ a utility function $U^c(S) : C \times S \rightarrow \mathbb{R}^+$ that measures the appropriateness of recommending an item s to user c . It is assumed that this function is not known for the whole $C \times S$ space

but only on some subset of it. Therefore, in the context of recommendation, we want for each user $c \in C$ to be able to:

- i estimate (or approach) the utility function $U^c(S)$ for an item s of the space S for which $U^c(S)$ is not yet known; or,
- ii choose a set of N items $s \in S$ that will maximise $U^c(S) : \forall c \in C$,
 $s = \operatorname{argmax}_{s \in S} U^c$

In most recommender systems, the utility function $U^c(S)$ usually considers one attribute of an item, e.g. its overall evaluation or rating. Nevertheless, utility may also involve more than one attribute of an item. The recommendation problem therefore becomes a multi-attribute one. We want to explore how the TEL recommendation problem can be better defined if such a multi-attribute modelling approach is followed in order to identify (Roy 1996):

- *Object of the decision*. That is, defining the object upon which the decision has to be made and the rationale of the recommendation decision.
- *Family of criteria*. That is, the identification and modelling of a set of criteria that affect the recommendation decision, and which are exhaustive and non-redundant.
- *Global preference model*. That is, the definition of the function that aggregates the marginal preferences upon each criterion into the global preference of the decision maker about each item.
- *Decision support process*. That is, the study of the various categories and types of recommender systems that may be used to support the recommendation decision maker, in accordance to the results of the previous steps.

In TEL recommendation, the object of decision is an item s that belongs to the set of all candidate items S representing any type of items that may be recommended to a user, such as a learning resource, a learning activity, a peer learner or a mentor. To express the rationale behind the decision, Roy (1996) refers to the notion of the decision *problematic*. The four types of common decision problematics identified in the Multi-Criteria Decision Making (MCDM) literature, may be considered valid in the context of TEL recommendation (Adomavicius et al. 2011):

- *Choice*, which involves choosing one item from a set of candidates;
- *Sorting*, which involves classifying items into pre-defined categories;
- *Ranking*, which involves ranking items from the best one to the worst one; and
- *Description*, which involves describing all the items in terms of performance upon each criterion.

For instance, in TEL contexts, a set of candidates may be learning resources, peer learners or learning activities. An example family of criteria that affects the recommendation decision can include the age, language, knowledge level, goal or other contextual variables such as the current device and available time of the learner. *Choice* involves choosing those objects that are appropriate for the learning setting and characteristics of the learner—such as learning resources to study the theory of relativity for K-12 learners in French. A TEL recommender system that supports *sorting* classifies items into predefined categories—in its simplest form such sorting

may constitute classifying items according to certain attributes, such as language or knowledge level. *Ranking* is the process of presenting items according to descending order of relevance. For instance, in a TEL context, such ranking may involve presenting those items that are most relevant for the age, knowledge level or mother tongue of the learner first. Finally, *description* involves presenting and explaining each one of the candidate items by analysing its predicted performance upon each one of the criteria. Such a description enables the user to gain insight into the various alternatives and help her take better informed decisions.

After the correct problematic is defined, the set of all candidate items S is analysed in terms of multiple criteria, in order to model all possible impacts, consequences, or attributes (Roy 1996). In recommender systems, the criteria may refer to multiple characteristics of an item (usually the case in content-based recommendation) or to the multiple dimensions upon which the item is being evaluated (the case in collaborative filtering recommendation). This step must conclude to a consistent family of n criteria $\{g_1, g_2, \dots, g_n\}$. In MCDM, four types of criteria are formally used (Jacquet-Lagreze and Siskos 2001):

- *Measurable*, is a criterion that allows quantified measurement upon an evaluation scale.
- *Ordinal*, is a criterion that defines an ordered set in the form of a qualitative or a descriptive scale.
- *Probabilistic*, is a criterion that uses probability distributions to cover uncertainty in the evaluation of alternatives.
- *Fuzzy*, is a criterion where evaluation of alternatives is represented in relationship to its possibility to belong in one of the intervals of the evaluation scale.

For instance, a measurable criterion that is used by a content-based recommender can be the age range of learners for which a learning resource is suitable. Examples of ordinal criteria are attributes that belong to an ordered controlled vocabulary, like the aggregation level of a resource and its easiness to understand. Probabilistic and fuzzy criteria are often used when a system has to deal with uncertainty of a criterion—such as an estimate of the knowledge level of the user or an estimate of the usefulness of a resource for a particular learning activity.

After the definition of the criteria, the development of a global preference model is made to provide a way to aggregate the values of each criterion g_i (with $i = 1, \dots, n$) in order to express the preferences between the different alternatives of the item set S . Examples of preference models include (Adomavicius et al. 2011):

- *Value-Focused models*, where a value system for aggregating the user preferences on the different criteria is constructed. In such approaches, marginal preferences upon each criterion are synthesised into a total value using a synthesising utility function (Keeney 1992).
- *Outranking Relations models*, where preferences are expressed as a system of outranking relations between two alternatives a and b , thus allowing the expression of incomparability. In such approaches, all alternatives are one-to-one compared between them, and preference relations are provided as relations *a is preferred*

to b , a is equally preferred to b , and a is incomparable to b (Roy and Bouyssou 1993).

- *Multi-Objective Optimisation models*, where criteria are expressed in the form of multiple constraints of a multi-objective optimisation problem. In such approaches, usually the goal is to find a Pareto optimal solution for the original optimisation problem (Zeleny 1974).
- *Preference Disaggregation models*, where the preference model is derived by analysing past decisions. Such approaches build on the models proposed by the previous ones (thus they are sometimes considered as a sub-category of other modelling approaches' categories), since they try to infer a preference model of a given form (e.g. value function) from some given preferential structures that have led to particular decisions in the past, and aim at producing decisions that are at least identical to the examined past ones (Jacquet-Lagrez and Siskos 2001).

The most typical cases of TEL recommender systems are value-focused (Manouselis et al. 2011), usually engaging a single-attribute (and rarely multi-attribute), linear, additive value function for the representation of user preferences. This is a traditional decision making approach, widely applied and convenient to implement. On the other hand, assuming that the preference function is single-attribute and linear restricts the way user preferences are represented. Therefore, alternative forms for representing preferences in a MCDM manner should be explored in TEL as well (Manouselis 2008; Adomavicius et al. 2011). This requirement is particularly relevant for TEL, where certain attributes may have a different influence on recommendation in different settings. For instance, both formal and informal learning processes have different requirements for the learning environment and, as such, for the recommendation within the environment. Often, it is not possible to draw a clear line between formal and informal learning scenarios. As an example, recommender systems need to deal with the tension of recommendations for activities liked by the learner and those required by the teacher (Tang and McCalla 2003). Since recommendations may differ depending on the context of the learner, it is therefore important to study carefully the intended recommendation goals to be supported. We identify such goals in the next section.

2.1.2 Identifying the TEL Recommendation Goals

In the past, the development of recommender systems has been related to a number of relevant user tasks that the recommender system supports within some particular application context. More specifically, Herlocker et al. (2004) have related popular (or less popular) user tasks with a number of specific recommendation goals that are included in Table 2.1. Generally speaking, most of these already identified recommendation goals and user tasks are valid in the case of TEL recommender systems as well. For example, a recommender system supporting learners to achieve a specific learning goal, providing annotation in context or recommending a sequence of

learning resources are relevant tasks. In the table, an example of how recommendation could support a similar user task is included for all the tasks that Herlocker et al. (2004) have identified. In addition, it includes a comment about any additional requirements that this brings forward for the developers of TEL recommender systems.

On the other hand, in comparison to the typical item recommendation scenario, there are several particularities to be considered regarding what kind of learning is desired, e.g. learning a new concept or reinforce existing knowledge may require different types of learning resources. This is reflected in the second part of Table 2.1, where examples of user tasks that are particularly interesting for TEL are in, extending the ones initially identified in Manouselis et al. (2011). Again, a comment on any additional requirements for developers of TEL recommenders is included.

Apart from this initial identification of tasks, recommendation in a TEL context has many particularities that are based on the richness of the pedagogical theories and models. For instance, for learners with no prior knowledge in a specific domain, relevant pedagogical rules such as Vygotsky's *zone of proximal development* could be applied: e.g. 'recommended learning objects should have a level slightly above learners' current competence level' (Vygotsky 1978). Different from buying products, learning is an effort that often takes more time and interactions compared to a commercial transaction. Learners rarely achieve a final end state after a fixed time. Instead of buying a product and then owning it, learners achieve different levels of competences that have various levels in different domains. In such scenarios, what is important is identifying the relevant learning goals and supporting learners in achieving them. On the other hand, depending on the context, some particular user task may be prioritised. This could call for recommendations whose time span is longer than the one of product recommendations, or recommendations of similar learning resources, since recapitulation and reiteration are central tasks of the learning process (McCalla 2004).

As for teacher-centered learning contexts, different tasks need to be supported. These tasks cover both the ones related to the preparation of lessons, the delivery of a lesson (i.e. the actual teaching), and the ones related to the evaluation/assessment. For instance, to prepare a lesson the teacher has certain educational goals to fulfill and needs to match the delivery methods to the profile of the learners (e.g. their previous knowledge). Lesson preparation can include a variety of information seeking tasks, such as finding content to motivate the learners, to recall existing knowledge, to illustrate, visualise and represent new concepts and information. The delivery can be supported in using different pedagogical methods (either supported with TEL or not), whose effectiveness is evaluated according to the goals set. A TEL recommender system could support one or more of these tasks, leading to a variety of recommendation goals.

Thus, although the previously identified user tasks and recommendation goals can be considered valid in a TEL context, there are several particularities and complexities. This means that simply transferring a recommender system from an existing (e.g. commercial) content to TEL may not accurately meet the needs of the targeted users. In TEL, careful analysis of the targeted users and their supported tasks should

Table 2.1 User tasks supported by current recommender systems and requirements for TEL recommender systems

Tasks	Description	Generic recommender	TEL recommenders	New requirements
Existing user tasks supported by recommender systems				
1. ANNOTATION IN CONTEXT	Recommendations while user carries out other tasks	E.g. predicting how relevant the links are within a web page	E.g. predicting relevance/usefulness of items in the reading list of a Moodle course or a Learning Network	Explore attributes for representing relevance/usefulness in a learning context
2. FIND GOOD ITEMS	Recommendations of suggested items	E.g. receiving list of web pages to visit	E.g. receiving a selected list of online educational resources around a topic	None
3. FIND ALL GOOD ITEMS	Recommendation of all relevant items	E.g. receiving a complete list of references on a topic	E.g. suggesting a complete list of scientific literature or blog postings around a topic	None
4. RECOMMEND SEQUENCE	Recommendation of a sequence of items	E.g. receive a proposed sequence of songs	E.g. receiving a proposed sequence through resources to achieve a particular learning goal	Explore formal and informal attributes for representing relevancy to a particular learning goal
5. JUST BROWSING	Recommendations out of the box while user is browsing	E.g. people that bought this, have also bought that	E.g. receiving recommendations for new courses on the university site or getting suggestions for additional blog postings in a Learning Network	Explore formal and informal attributes for representing relevance/usefulness in a learning context
6. FIND CREDIBLE RECOMMENDER	Recommendations during initial exploration/testing phase of a system	E.g. movies that you will definitely like	E.g. restricting initial recommendations to ones with high confidence /credibility	Explore criteria for measuring confidence and credibility in formal and informal learning

(Continued)

Table 2.1 (continued)

Tasks	Description	Generic recommender	TEL recommenders	New requirements
TEL user tasks that could be supported by recommender systems				
1. FIND NOVEL RESOURCES	Recommendations of particularly new or novel items	E.g. receiving recommendations about latest additions or particularly controversial items	E.g. receiving very new and/or controversial resources on covered topics	Explore recommendation techniques that select items beyond their similarity
2. FIND PEERS	Recommendation of other people with relevant interests	E.g. being suggested profiles of users with similar interests	E.g. being suggested student in the same class or a peer-student in a Learning Network	Explore attributes for measuring the similarity with other people
3. FIND GOOD PATHWAYS	Recommendation of alternative learning paths through learning resources	E.g. receive alternative sequences of similar songs	E.g. receiving a list of alternative learning paths over the same resources to achieve a learning goal depending	Explore criteria for the construction and suggestion of alternative (but similar) sequences
4. PREDICT STUDENT PERFORMANCE	Prediction of student performance based on previous behavior	E.g. predicting student performance in visual graphs compared to average student scores	E.g. recommending group combinations for teachers to improve course performance, recommending learning activities to improve individual student performance	Take advantage of student data from LMS, tracking of students and teacher activities, critical interpretation of analysed data

be carried out, before a recommendation goal is defined and a recommender system is deployed. This means that the TEL recommendation goals can be rather complex: for example, a typical TEL recommender system could suggest a number of alternative learning paths throughout a variety of learning resources, either in the form of learning sequences or hierarchies of interacting learning resources. This should take place in a pedagogically meaningful way that will reflect the individual learning goals and targeted competence levels of the user, depending on proficiency levels, specific interests and the intended application context. A number of context variables have to be considered, such as user attributes, domain characteristics, and intelligent methods that can be engaged to provide personalised recommendations.

2.1.3 Identifying the TEL Context Variables

As outlined by Romero and Ventura (2007), the TEL domain differs from domains like e-commerce in several ways. In e-commerce, the used data are often simple web server access logs or ratings of users on items. In TEL, many researchers use more information about a learner interaction (Pahl and Donnellan 2002). The user model and the objectives of the systems are also different in both application domains (Drachsler et al. 2009a).

A survey of existing TEL interaction data models has been presented in Butoianu et al. (2010). Examples of models to represent learner interactions are the User Interaction Context Ontology (UICO, Rath et al. 2009) and Contextualised Attention Metadata (CAM, Scheffel et al. 2011) models. Both models capture actions of the user, such as *select*, *save*, *create* and *write* actions, on resources. In addition, the context in which an action occurred, such as the current task of the learner, can be captured. The Atom activity stream Resource Description Framework (RDF) mapping of the LinkedEducation.org initiative presents a similar approach to model actions of users in social networks. Vocabularies for actions, actors and objects involved and related contextual information are defined.

In addition to interaction models, researchers in TEL have elaborated learner models that describe several characteristics of learners. Brusilovsky and Millan (2007) identified the following categories based on an analysis of the existing literature: *knowledge levels*, *goals and tasks*, *interests*, *background* and *learning and cognitive styles*. In addition, several models, standards and specifications have been elaborated to describe learning resources. The IEEE LOM and Dublin Core metadata standards are prominently used by TEL applications to describe characteristics of learning resources, including general characteristics, such as title, author and keywords, technical and educational characteristics and relations between learning resources.

We integrated the various data categories and elements in Fig. 2.1. We use this framework in the remainder of this chapter to identify data elements in existing data sets. The model has been developed by synthesising existing works on interaction data and context variables in the TEL field that were outlined above. It could be further refined by studying relevant theoretical frameworks, like the Activity

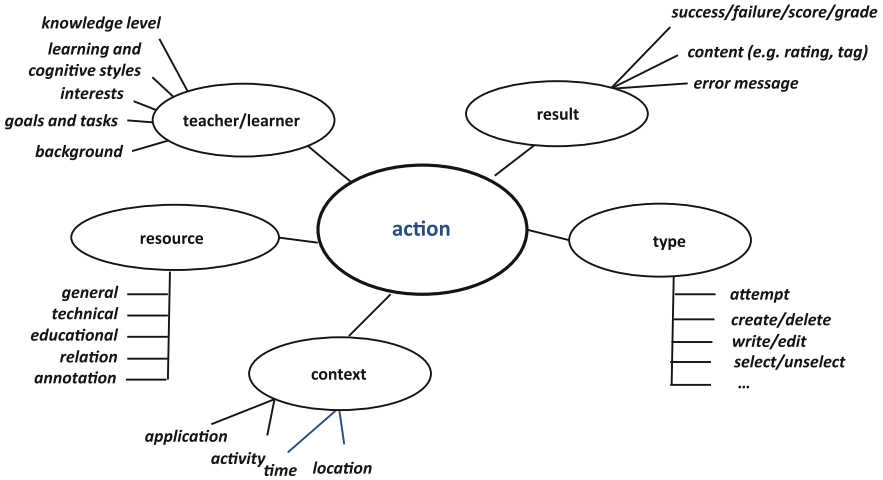


Fig. 2.1 TEL variables (adapted from Verbert et al. (to appear))

Theory (Kaptelinin et al. 1995), that could help reorganise the various categories and elements.

2.2 Data Sets to Support TEL Recommendation

2.2.1 Collecting TEL Data Sets

An important requirement to facilitate research on recommendation technologies is the existence of sufficient data from various system activities and its interactions with users. When the analysis is taking place for research purposes and in an exploratory manner, it is equally important to provide researchers with sufficient data coming from a real or simulated environment of the targeted domain. In an increasing number of scientific disciplines, large data collections are emerging as important community resources (Chervenak et al. 2000). These data sets are used as benchmarks to develop new algorithms and compare them to other algorithms in given settings. In data sets that are used for recommendations algorithms, such data can for instance be explicit (ratings) or implicit (downloads and tags) relevance indicators. These indicators are then for instance used to find users with similar interests as a basis to suggest items to a user.

To collect TEL data sets, the first dataTEL Challenge was launched as part of the Workshop on Recommender Systems for TEL (Manouselis et al. 2010a), jointly organised by the 4th ACM Conference on Recommender Systems and the 5th European Conference on Technology Enhanced Learning in September 2010. In this call, research groups were invited to submit existing data sets from TEL applications.

As a follow up activity, the dataTEL—Data sets for Technology Enhanced Learning Workshop was organised at the Second STELLAR Alpine Rendez-Vous in March 2011 (Drachsler et al. 2011, to appear). During this workshop, researchers discussed related initiatives that are collecting educational data sets, additional data sets that are relevant for recommendation research, as well as challenges related to privacy and data protection and research on evaluation methodologies.

Similar work is carried out at the Pittsburgh Science of Learning Center (PSLC). The PSLC DataShop (Stamper et al. 2010) is a data repository that provides access to a large number of educational data sets derived from intelligent tutoring systems. Currently, 270 data sets are stored that record 58 million learner actions. Several researchers of the educational data mining community have used these data sets.

The Mulce project (Reffay and Betbeder 2009) is also collecting and sharing interaction data of learners. A platform is available to share, browse and analyse shared data sets. At the time of writing, 34 data sets are available on the portal, including a data set of the Virtual Math Teams project that investigated the use of online collaborative environments for K-12 mathematics learning. These data sets have been used extensively by the Computer Supported Collaborative Learning (CSCL) community.

LinkedEducation.org is another initiative that provides an open platform to promote the use of data for educational purposes. Available data sets describe the structure of organisations and institutions, the structure of courses, learning resources and interrelationships between people. Schemas and vocabularies are provided to describe discourse relationships and activity streams. Such schemas and vocabularies offer interesting perspectives for the sharing and reuse of interaction data between users that is relevant for research on recommendation techniques.

Several other initiatives are available that focus on providing the means to share data sets among researchers in a more generic way. DataCite.org is an organisation that enables users to register research data sets and to assign persistent identifiers to them, so that data sets can be handled as citable scientific objects. The Dataverse Network (King 2007) is an open-source application for publishing, citing and discovering research data. Fact sheets of data sets are gathered from organisations and researchers are encouraged to make data publicly available.

In this section, we analyse educational data sets that have been collected by dataTEL (Drachsler et al. 2010a) and that can be used for research on recommendation for learning. A detailed description of other data sets and their usefulness for a wider variety of research purposes may be found at Verbert et al. (2011) and Verbert et al. (to appear).

2.2.2 Collected Data Sets

Table 2.2 presents the data sets have been collected as a result of the first dataTEL challenge:

Table 2.2 Overview data sets

	Mendeley	APOSdle	ReMashed	Organic Edunet	Mace	Travel well	CGIAR
Collection period	1 year	3 months	2 years	9 months	3 years	6 months	6 years
Number of users	200.000	6	140	1.000	1.148	98	841
Number of items	1.857.912	163	96.000	10.500	12.000	1.923	14.693
Number of actions	4.848.725	1.500	23.264	920	461.982	16.353	326.339
Publicly available	+	+	-	-	-	+	-
<i>Action</i>							
Attempt	-	-	-	-	-	-	+
Select/unselect	+	+	-	-	+	-	+
Tags	-	+	+	+	+	+	-
Rate/star	+	-	+	+	+	+	-
Save/download	+	+	-	-	-	-	-
Search	-	+	-	-	-	-	+
<i>Learner/teacher</i>							
Id	+	+	+	+	+	+	+
Knowledge level	-	+	+	-	-	-	-
Interests	-	+	+	-	-	+	-
<i>Resource</i>							
General	+	+	+	+	+	+	+
Technical	-	-	-	-	+	+	-
Educational	-	-	-	-	+	+	-
Annotation	-	-	+	+	+	+	-
<i>Context</i>							
Time	-	+	+	+	+	+	+
Activity	-	+	-	-	-	-	-
<i>Result</i>							
Success/failure	-	+	-	-	-	-	+
Content (e.g. rating, tag)	-	-	+	+	+	+	+

- The first data set was submitted by *Mendeley* (Jack et al. [to appear](#)), a research platform that helps users organise research papers and collaborate with colleagues. In the context of learning, such a data set provides useful data for recommender systems that suggest papers to learners or teachers or to suggest suitable peer learners on the basis of common research or learning interests. The data set contains *learner* and *resource* data. For each learner, implicit interest data is available that captures which papers a user has selected, starred or added to her library.
- The *APOSDLE* data set originates from the APOSDLE (Ghidini et al. [2007](#)) project. APOSDLE is an adaptive learning system that aims to support learning within everyday work tasks. The system recommends resources (documents, videos, links) and colleagues who can help a user with a task. The data set captures information about the *activity context* (tasks, topics), the *learner* (knowledge level, implicit interest indicators) and actions on *resources*.
- The *ReMashed* data set focuses on community knowledge sharing (Drachsler et al. [2010b](#)). The data set includes information about interests of *learners* (ratings and tags) on available *resources*. The main objective of ReMashed is to offer personalized information access to the emergent information space of the community.
- The *MACE* data set originates from the MACE *eContentplus* project (Wolpers et al. [2009](#)). The MACE portal provides advanced graphical metadata based access to learning resources in architecture that are stored in different repositories all over Europe. The data set contains both implicit (search activities, select, downloads, tags) and explicit (ratings) interest data of the *learner* on *resources*. In addition, the *time* of each user activity is recorded.
- The *Organic.Edunet* data set was collected on the Organic.Edunet Web portal (Manouselis et al. [2009](#)), a learning portal for organic agriculture educators that provides access a large number of learning resources from a federation of 11 institutional repositories. The data set includes *learner* (interest data in the form of tags and ratings) and *resource* data. The particularity of this data set is the fact that ratings are collected upon three different criteria: the usefulness of a resource as a learning tool, the relevance to the organic thematic, and the quality of its metadata.
- The *Travel well* data set originates from the MELT *eContentplus* project (Vuorikari [2009](#)). The data set was collected on the MELT Learning Resource Exchange portal that makes open educational resources available from 20 content providers in Europe and elsewhere. This data set includes information about *teachers* (interest indicators in the form of tags and ratings and topics of interests as provided by the teachers), *resources* (minimum age, maximum age, duration, resource type, duration) and the *timestamp* of user actions.
- The *CGIAR* contains data from a Moodle installation used for agroforestry courses of CGIAR¹—a worldwide network of Agricultural Research Centers. The data set contains actions of learners on quizzes and retrieval of documents provided by the teacher through a Moodle LMS.

¹ <http://cgiar.org>

The CGIAR, MACE, Organic.Edunet and Mendeley data sets are the largest data sets that contain data of 841, 1.148, 1000 and 200.000 users during a time period of 6 years, 3 years, 9 months and 1 years, respectively. The Travel well data set contains ratings and tags of about 100 users that were collected during a six month time period. The ReMashed data sets collects activities of 140 users during a 2 year period. The current APOSDLE data set is only a sample that captures data of a few users only.

Several data sets have been collected that are openly accessible. Registration is sometimes required before a dataset can be downloaded. For other datasets, legal protection rules apply. We obtained these data sets by sending a statement of our intended research purposes to the organisation. These statements were then in most cases analysed by their legal department before approval was granted. All data sets contain data that is anonymised, so that it can no longer be linked to an individual.

2.2.3 Usefulness for TEL Recommender Systems

Several dataTEL data sets contain relevance indicators that are useful for research on recommendation algorithms for learning. Of interest in this discussion are the data elements that are provided by the data sets. Explicit relevance feedback, such as ratings by users, are provided in the MACE, ReMashed, Organic.Edunet and Travel well data sets. These data sets provide ratings on a five point likert scale and are interesting data sets for developing recommendation algorithms to *find novel resources*. Mendeley provides information on articles that are starred by a user ('1' if the article has been starred and '0' otherwise), but the semantics of such stars in user libraries may be different for different users (i.e. a star can indicate relevance feedback, but may as well indicate that the user wants to read the article at a later stage). Therefore, the application of such data for recommendation is less straightforward. In addition, implicit relevance indicators, such as downloads, search terms and annotations, are available. If time interval data is available, the data might be suitable to extract reading times in order to determine the relevancy of a resource.

Manouselis et al. (2007, 2010b) used the Travel well data set to evaluate recommendation algorithms for learning. Similar experiments have been reported in Verbert et al. (2011). In this study, the Mendeley and MACE data sets were also used. Although still preliminary, some conclusions were drawn about successful parameterisation of collaborative filtering algorithms for learning. Outcomes suggest that the use of implicit relevance indicators, such as downloads, tags and read actions, are useful to suggest learning resources.

The data sets can also be used for research to *find peers*. For instance, by analysing interaction patterns of learners, a recommender system may identify peer helpers who are able to help with a learning activity. In addition, data sets derived from web portals, such as the Organic.Edunet, MACE, Mendeley, and Travel well data sets, can be used for finding users with common interests. Such prediction of user attributes has been researched extensively by the Educational Data Mining community and includes finding estimates of the knowledge level of a user based on interaction data.

An extensive overview of research in this area has been presented in Romero and Ventura (2007).

Finding good pathways is a third recommendation task that is relevant for a TEL context. There are several ways to support research on recommendation of such learning sequences. Time information can be used to extract sequencing patterns from data sets that capture interactions of users with resources, such as select, annotate, rate or download actions. Such information is available in many dataTEL data sets, including the ReMashed, MACE, Organic.Edunet, Travel well and CGIAR data sets. Drachsler et al. (2009b) researched the influence of sequence recommendation on the learning process with the ReMashed data set. An alternative way to support sequence recommendation would be to find pathways based on other learner characteristics, such as knowledge level. Such research has been conducted by the Intelligent Tutoring Systems community. Cheung et al. (2003) for instance suggest next learning resources that relate to prior knowledge of the course in order to provide good “orientation”, which is the pathway to learn the material.

Predicting student performance is a fourth recommendation task that has been researched extensively over the last decade. Several data sets are available that can support research on prediction of learner performance and discovery of learner models. Among others, such predictions are researched to provide advice when a learner is solving a problem (Romero and Ventura 2007). Data sets from intelligent tutoring systems that capture attempts of learners provide a rich source of data to estimate the knowledge level of a learner. Some data sets derived from LMSs, such as the CGIAR data set, contain data on the number of attempts and total time spent on assignments, forums and quizzes. Romero et al. (2008) compared different data mining techniques to classify learners based on such LMS data and the final grade obtained for courses. A more extensive analysis of usefulness of data sets for learning analytics purposes has been presented in Verbert et al. (to appear).

References

- G. Adomavicius, A. Tuzhilin, Towards the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Trans. Knowl. Data Eng.* **17**(6), 734–749 (2005)
- G. Adomavicius, N. Manouselis, Y. Kwon, Multi-criteria recommender systems. in *Recommender Systems Handbook*, ed. by F. Ricci, L. Rokach, B. Shapira (eds.) (Springer US, 2011), pp. 769–803
- P. Brusilovsky, E. Millan, User models for adaptive hypermedia and adaptive educational systems, in *The adaptive web*, vol. 4321, ed. by P. Brusilovsky, et al. (Springer, 2007), pp. 3–53
- V. Butoianu, P. Vidal, K. Verbert, E. Duval, J. Broisin, User context and personalized learning: a federation of contextualized attention metadata. *J. UCS* **16**(16), 2252–2271 (2010)
- A. Chervenak, I. Foster, C. Kesselman, C. Salisbury, S. Tuecke, The data grid: towards an architecture for the distributed management and analysis of large scientific data sets. *J. Netw. Comput. Appl.* **23**(3), 187–200 (2000)
- B. Cheung, L. Hui, J. Zhang, S.M. Yiu, SmartTutor: an intelligent tutoring system in web-based adult education. *J. Syst. Softw.* **68**(1), 11–25 (2003)
- M. Deshpande, G. Karypis, Item-based Top-N recommendation algorithms. *ACM Trans. Inf. Syst.* **22**(1), 143–177 (2004)

- H. Drachsler, H.G.K. Hummel, R. Koper, Identifying the goal, user model and conditions of recommender systems for formal and informal learning. *J. Digit. Inf.* **10**(2), 4–24 (2009a)
- H. Drachsler, H. Hummel, B. Berg, J. Eshuis, W. Waterink, R. Nadolski, A. Berlanga, N. Boers, R. Koper, Evaluating the effectiveness of personalised recommender systems in learning networks. in *Learning Network Services for Professional Development*, ed. by R. Koper (eds.) (Springer, Berlin Heidelberg, 2009b), pp. 95–113
- H. Drachsler, T. Bogers, R. Vuorikari, K. Verbert, E. Duval, N. Manouselis, G. Beham, S. Lindstaedt, H. Stern, M. Friedrich, Issues and considerations regarding sharable data sets for recommender systems in technology enhanced learning. *Procedia Comput. Sci.* **1**(2), 2849–2858 (2010a). doi:[10.1016/j.procs.2010.08.010](https://doi.org/10.1016/j.procs.2010.08.010)
- H. Drachsler, K. Verbert, M.-A. Sicilia, M. Wolpers, N. Manouselis, R. Vuorikari, S. Lindstaedt, F. Fischer, DataTEL—datasets for technology enhanced learning. STELLAR ARV White Paper (2011), http://oa.stellarnet.eu/open-archive/browse?resource=6756_v1. Accessed 12 March 2012
- H. Drachsler, L. Rutledge, P. van Rosmalen, H. Hummel, D. Pecceu, T. Arts, E. Hutten, R. Koper, Remashed—an usability study of a recommender system for mash-ups for learning. *International Journal of Emerging Technologies in Learning (iJet)*, Special Issue: ICL2009 MashUps for Learning **5**, 7–11 (2010b)
- H. Drachsler, K. Verbert, N. Manouselis, S. Lindstaedt, R. Vuorikari, M. Wolpers, Preface for data-TEL Special Issue on Datasets and Data Supported Learning in Technology-Enhanced Learning. Drachsler H, Verbert K, Manouselis N, Lindstaedt S, Vuorikari R, Wolpers M (eds) *International Journal of Technology Enhanced Learning (IJTEL)*, (to appear)
- C. Ghidini, V. Pammer, P. Scheir, L. Serafini, S. Lindstaedt, Aposdle: learn@ work with semantic web technology. *Proceedings of ISEMANTICS*, pp. 262–269 (2007)
- J.L. Herlocker, J.A. Konstan, L.G. Terveen, J.T. Riedl, Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst.* **22**(1), 5–53 (2004)
- K. Jack, M. Hristakeva, R. Garcia de Zuniga, M. Granitzer, Mendeley’s Open Data for Science and Learning: A Reply to the DataTEL Challenge. *International Journal of Technology Enhanced Learning* (to appear)
- E. Jacquet-Lagrez, Y. Siskos, Preference disaggregation: 20 years of MCDA experience. *Eur. J. Oper. Res.* **130**, 233–245 (2001)
- V. Kaptelinin, K. Kuutti, L.J. Bannon, Activity theory: basic concepts and applications, in *Selected papers from the 5th International Conference on Human-Computer Interaction*, ed. by B. Blumenthal, et al. (Springer-Verlag, London, 1995), pp. 89–201
- R.L. Keeney, *Value-focused Thinking: A Path to Creative Decisionmaking* (Harvard University Press, Cambridge MA, 1992)
- G. King, An introduction to the dataverse network as an infrastructure for data sharing. *Sociol. Methods Res.* **36**(2), 173–199 (2007)
- J.A. Konstan, Introduction to recommender systems: algorithms and evaluation. *ACM Trans. Inf. Syst.* **22**(1), 1–4 (2004)
- N. Manouselis, H. Drachsler, K. Verbert, O.C. Santos, RecSysTEL preface. *Procedia Comput. Sci.* **1**(2), 2773–2774 (2010a)
- N. Manouselis, R. Vuorikari, F. Van Assche, Collaborative recommendation of e-learning resources: an experimental investigation. *JCAL* **26**(4), 227–242 (2010b)
- N. Manouselis, (2008) A discussion on multi-criteria recommendation. in *Proceedings of the Workshop on Recommender Systems, 18th European Conference on Artificial Intelligence (ECAI 2008)*, Patras, Greece, July, 2008
- N. Manouselis, R. Vuorikari, F. Van Assche, Simulated analysis of MAUT collaborative filtering for learning object recommendation. in *Proceedings of the Workshop on Social Information Retrieval in Technology Enhanced Learning (SIRTEL 2007)*, Crete, Greece, 2007
- N. Manouselis, H. Drachsler, R. Vuorikari, H. Hummel, R. Koper, Recommender Systems in Technology Enhanced Learning, in *Kantor P*, ed. by F. Ricci, L. Rokach, B. Shapira (Recommender Systems Handbook, Springer US, 2011), pp. 387–415

- N. Manouselis, K. Kastrantas, S.S. Alonso, J. Caceres, H. Ebner, M. Palmer, Architecture of the organic.Edunet web portal. *Int. J. Web Portals* **1**(1), 71–91 (2009)
- G. McCalla, The ecological approach to the design of E-learning environments: purpose-based capture and use of information about learners. *Journal of Interactive Media in Education*, Special Issue on the Educational Semantic Web, 7, ISSN:1365–893X (2004)
- C. Pahl, C. Donnellan, Data mining technology for the evaluation of web-based teaching and learning systems, in *Proceedings of E-Learn 2002*, ed. by M. Driscoll, T. Reeves (AACE, Chesapeake, VA, 2002), pp. 747–752
- A. Rath, D. Devaurs, S. Lindstaedt, UICO: an ontology-based user interaction context model for automatic task detection on the computer desktop, in *Proceedings of the 1st Workshop on Context, Information and Ontologies*, p. 10, ACM (2009)
- C. Reffay, M.-L. Betbeder, Sharing corpora and tools to improve interaction analysis, in *Proceedings of EC-TEL '09, LNCS*, vol. 5794, ed. by U. Cress, V. Dimitrova, M. Specht (Springer-Verlag, Berlin, Heidelberg, 2009), pp. 196–210
- C. Romero, S. Ventura, Educational data mining: a survey from 1995 to 2005. *Expert Syst. Appl.* **33**(1), 135–146 (2007). Elsevier
- C. Romero, S. Ventura, P.G. Espejo, C. Hervs, Data mining algorithms to classify students, in *Proceedings of the 1st International Conference on Educational Data Mining*, ed. by R. de Baker, T. Barnes, J. Beck (eds) pp. 8–17 (2008)
- B. Roy, D. Bouyssou, *Aide Multicritère a la Decision: Methodes et Cas* (Economica, Paris, 1993)
- B. Roy, *Multicriteria Methodology for Decision Aiding* (Kluwer Academic Publishers, Dordrecht, 1996)
- M. Scheffel, K. Niemann, A. Pardo, D. Leony, M. Friedrich, K. Schmidt, M. Wolpers, C.D. Kloos, Usage pattern recognition in student activities, in *Proceedings of EC-TEL'11* (Springer-Verlag, Berlin, Heidelberg, 2011), pp. 341–355
- J.C. Stamper, K.R. Koedinger, R. Baker, A. Skogsholm, B. Leber, J. Rankin, S. Demi, PSLC dataShop: a data analysis service for the learning science community, in *Proceedings of Intelligent Tutoring Systems*, vol. 6095, (Springer, 2010), pp. 455–456
- T. Tang, G. McCalla, Smart recommendation for an evolving E-learning system, in *Proceedings of the Workshop on Technologies for Electronic Documents for Supporting Learning, International Conference on Artificial Intelligence in Education (AIED 2003)*, (2003) pp. 699–710
- K. Verbert, H. Drachsler, N. Manouselis, M. Wolpers, R. Vuorikari, E. Duval, Dataset-driven research for improving recommender systems for learning, in *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, pp. 44–53. ACM, New York, NY, USA. ISBN: 978-1-4503-0944-8, 2011
- K. Verbert, N. Manouselis, H. Drachsler, E. Duval, Dataset-driven Research to Support Learning and Knowledge Analytics. *Educational Technology and Society*, Special Issue on Learning and Knowledge Analytics, (to appear)
- R. Vuorikari, Ecology of social search for learning resources. *CampusWide Inf. Syst.* **26**(4), 272–286 (2009)
- L. Vygotsky, *Mind in Society: The Development of Higher Psychological Processes* (Harvard University Press, 1978)
- M. Wolpers, M. Memmel, A. Giretti, Metadata in architecture education—first evaluation results of the mace system. *Learn. Synerg. Multiple Discipl.* **5794**, 112–126 (2009)
- M. Zeleny, *Linear Multiobjective Programming* (Springer, New York, 1974)

<http://www.springer.com/978-1-4614-4360-5>

Recommender Systems for Learning

Manouselis, N.; Drachsler, H.; Verbert, K.; Duval, E.

2013, XI, 76 p. 4 illus., Softcover

ISBN: 978-1-4614-4360-5