dataTEL
Data-driven Research and Learning Analytics

EATEL-Special Interest Group

Hendrik Drachsler (a), Katrien Verbert (b)

(a) CELSTEC, Open University of the Netherlands
(b) Dept. Computer Science, K.U.Leuven, Belgium
Who is dataTEL?
dataTEL is a Theme Team funded by the STELLAR network of excellence

Riina Vuorikari
Stephanie Lindstaedt
Katrien Verbert
Nikos Manouselis
Martin Wolpers
Hendrik Drachsler
Who is dataTEL?
dataTEL is a Theme Team funded by the STELLAR network of excellence

Riina Vuorikari
Stephanie Lindstaedt
Katrien Verbert
Nikos Manouselis
Martin Wolpers
Hendrik Drachsler

MAVSEL

CEN PT
Social Data

Miguel Angel Sicillia
Joris Klerkx
Chapter 12
Recommender Systems in Technology Enhanced Learning

Nikos Manouselis, Hendrik Drachsler, Rina Voskamp, Haas Hurnael and Rah Koper

Abstract Technology enhanced learning (TEL) aims to design, develop and test socio-technical innovations that will support and enhance learning practices of both individuals and organisations. It is therefore an application domain that generally covers technologies that support all forms of teaching and learning activities. Since information retrieval (in terms of searching for relevant learning resources to support teachers or learners) is a pivotal activity in TEL, the deployment of recommender systems has attracted increased interest. This chapter attempts to provide an introduction to recommender systems for TEL settings, as well as to highlight their particularities compared to recommender systems for other application domains.

Nikos Manouselis
Greek Research and Technology Network (GRNET S.A.), 56 Mesogion Ave., 115 27, Athens, Greece e-mail: nikos@grnet.gr

Hendrik Drachsler
Centre for Learning Sciences and Technologies (CELSITEC), Open Universiteit Nederland e-mail: hendrik.drachsler@oog.nl

Rina Voskamp
European Schoolnet (EUN), 24, Rue Paul Émile Janson, 1050 Brussels, Belgium e-mail: r.voskamp@eun.org
Chapter 12
Recommender Systems in Technology Enhanced Learning

Nikos Manouselis, Hendrik Drachsler, Rina Vuoriari, Haas Hummel and Rah Koper

Abstract Technology enhanced learning (TEL) aims to design, develop and test socio-technical innovations that will support and enhance learning practices of both individuals and organisations. It is therefore an application domain that generally covers technologies that support all forms of teaching and learning activities. Since information retrieval (in terms of searching for relevant learning resources to support learning) is a common requirement in TEL systems, recommender systems can be considered as a promising way to support these tasks. This chapter provides an overview of recommender systems applications in the TEL domain. It presents a state-of-the-art view on the recurring challenges and peculiarities of the TEL domain. Furthermore, a number of key research questions are identified to help advance the design of recommender systems for TEL.

Table 12.3: Implemented TEL recommender systems reported in literature

<table>
<thead>
<tr>
<th>System</th>
<th>Status</th>
<th>Evaluator focus</th>
<th>Evaluation roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altered Vista [81, 82, 82, 104]</td>
<td>Full system</td>
<td>Interface, Algorithm, System usage</td>
<td>Human users</td>
</tr>
<tr>
<td>RACOFI [2, 61]</td>
<td>Prototype</td>
<td>Algorithm</td>
<td>System designers</td>
</tr>
<tr>
<td>QSAI [78, 79]</td>
<td>Full system</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>CYCLADES [4]</td>
<td>Full system</td>
<td>Algorithm</td>
<td>System designers</td>
</tr>
<tr>
<td>CoFind [29, 30]</td>
<td>Prototype</td>
<td>System usage</td>
<td>Human users</td>
</tr>
</tbody>
</table>
The TEL recommender research is a bit like this...
But...

“The performance results of different research efforts in TEL recommender systems are hardly comparable.”

(Manouselis et al., 2010)
But...

The TEL recommender experiments lack transparency. They need to be repeatable to test:

• Validity
• Verification
• Compare results
1st Workshop on Recommender Systems for Technology Enhanced Learning (RecSysTEL 2010)

Issues and Considerations regarding Sharable Data Sets for Recommender Systems in Technology Enhanced Learning

Hendrik Drachsler*, Toine Bogers+, Riina Vuorikari+, Katrien Verbert+, Erik Duval+, Nikos Manouselis+, Guenter Beham+, Stephanie Lindstaedt+, Hermann Stern+, Martin Friedrich+, Martin Wolpers+

RecSysTEL 2010
Workshop on Recommender Systems for Technology Enhanced Learning

Organised jointly by
4th ACM Conference on Recommender Systems (RecSys 2010)
5th European Conference on Technology Enhanced Learning (EC-TEL 2010)

Barcelona, Spain, 29-30 September 2010
## dataTEL::Collection

<table>
<thead>
<tr>
<th></th>
<th>Mendeley</th>
<th>APOSDE</th>
<th>ReMashed</th>
<th>Organic edunet</th>
<th>Mace</th>
<th>Melt</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Collection period</strong></td>
<td>1 year</td>
<td>3 months</td>
<td>2 years</td>
<td>9 months</td>
<td>3 years</td>
<td>6 months</td>
</tr>
<tr>
<td><strong>Users</strong></td>
<td>200,000</td>
<td>6</td>
<td>140</td>
<td>1,000</td>
<td>1,148</td>
<td>98</td>
</tr>
<tr>
<td><strong>Items</strong></td>
<td>1,857,912</td>
<td>163</td>
<td>96,000</td>
<td>11,000</td>
<td>12,000</td>
<td>1,923</td>
</tr>
<tr>
<td><strong>Activities</strong></td>
<td>4,848,725</td>
<td>1,500</td>
<td>23,264</td>
<td>920</td>
<td>461,982</td>
<td>16,353</td>
</tr>
<tr>
<td><strong>reads</strong></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td><strong>tags</strong></td>
<td>+</td>
<td>(+)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>ratings</strong></td>
<td>(+)</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>downloads</strong></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>search</strong></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td><strong>collaborations</strong></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>tasks/goals</strong></td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>sequence</strong></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Dataset-driven Research for Improving Recommender Systems for Learning

Katrien Verbist\textsuperscript{1}, Hendrik Drachsler\textsuperscript{2}, Nikos Manouselis\textsuperscript{3}, Martin Wolpers\textsuperscript{4}, Riina Vuorikari\textsuperscript{5} and Erik Duval\textsuperscript{1}

\textsuperscript{1} Department of Computer Science, K.U.Leuven
Celestijnenlaan 200A, B-3001 Leuven, Belgium
\{katrien.verbist, erik.duval\}@cs.kuleuven.be
\textsuperscript{2} Open University of the Netherlands (OUNL)
P.O. Box 2960, 6401 DL Haren, The Netherlands
hendrik.drachsler@ou.nl
\textsuperscript{3} Greek Research and Technology Network (GRNET)
56 Messogeion Av., 115 27, Athens, Greece
nikos@grnet.gr
\textsuperscript{4} Fraunhofer Institute for Applied Information Technology (FIT)
Schloss Birlinghoven, 53754 Sankt Augustin, Germany
martin.wolpers@fit.fraunhofer.de
\textsuperscript{5} European Schoolnet (EUN)
Rue de Trèves, 61, 1040 Brussels, Belgium
riina.vuorikari@eun.org

Abstract. In the world of recommender systems, it is a common practice to use public available datasets from different application environments (e.g. MovieLens, Book Crossing, or EachMovie) in order to evaluate recommendation algorithms. These datasets are used as benchmarks to develop new recommendation algorithms and to compare them to other algorithms in given settings. In this paper, we explore datasets that capture learner interactions with tools and resources. We use the datasets to evaluate and compare the performance of different recommendation algorithms for Technology Enhanced Learning (TEL). We present an experimental comparison of the accuracy of several collaborative filtering algorithms applied to these TEL datasets and elaborate on implicit relevance data, such as downloads and tags, that can be augment explicit relevance evidence in order to improve the performance of recommendation algorithms.
STELLAR Alpine Rendez-Vous White Paper

Workshop 6: dataTEL - Datasets for Technology Enhanced Learning

Workshop Organisers
Hendrik Drachsler, (Open University of the Netherlands, NL)
Katrien Verbert (K.U. Leuven, BE)
Miguel-Angel Sicilia (University of Alcalá, ES)
Martin Wolpers (Fraunhofer Institute for Applied Information Technology, DE)
Nikos Manouselis (Greek Research and Technology Network, GR)
Raija Vuorikari (European Schoolnet, BE)
Stefanie Lindstaedt (KnowCenter, AT)

Workshop Provocateur:
Frank Fischer (LMU Munich, DE)

powered by

[STEMellar and dataTEL logos]
Grand Challenges

1. **Topic**: Evaluation of recommender systems in TEL

2. **Topic**: Data supported learning examples

3. **Topic**: Datasets from learning object repositories and web content

4. **Topic**: Privacy and data protection for educational datasets
Dataset-driven Research to Support Learning and Knowledge Analytics

Katrien Verbert\(^1\), Nikos Manouselis\(^2\), Hendrik Drachsler\(^3\) and Erik Duval\(^2\)

\(^1\)Dept. Computerwetenschappen, K.U.Leuven, Belgium
\(^2\)Agro-Know, Athens, Greece & University of Alcala, Spain
\(^3\)CELSTEC, Open University of the Nederlands, The Nederlands

katrien.verbert@cs.kuleuven.be // nikosma@ieee.org // hendrik.drachsler@ou.nl
// erik.duval@cs.kuleuven.be

ABSTRACT

In various research areas, the availability of open data is considered as key for research and application purposes. For instance, in the world of e-commerce, it is a common practice to use publicly available datasets from different application environments (e.g., MovieLens or Bookcrossing) to support research on recommendation techniques. These datasets are used as benchmarks to develop new algorithms and to compare them to other algorithms in given settings. Finding such available datasets for experimentation can be a challenging task in technology enhanced learning, as there are various sources of data that have not been identified and documented exhaustively. In this paper, we provide such an analysis of datasets that can be used for research on learning and knowledge analytics. First, we present a framework for the analysis of educational datasets. Then, we analyze existing datasets along the dimensions of this framework and outline future challenges for the collection and sharing of educational datasets.

Keywords
Learning and knowledge analytics, Datasets, Open science

Introduction

The need for better measurement, collection, analysis and reporting of data about learners has been identified by several researchers in the Technology Enhanced Learning (TEL) field (Siemens 2010, Romero et al. 2007, Drachsler et al. 2010). This need has been translated into an emerging strand of research on learning and knowledge analytics (LAK), as reflected by a number of conferences and special issues in recent years (Siemens & Gasevic 2011). Among others, the analysis of learner data and identification of patterns within these data are researched to predict learning outcomes, to suggest relevant resources and to detect error patterns or affects.
Who is dataTEL?
dataTEL - EATEL SIG on Data-driven Research and Learning Analytics

In the educational world, only very limited datasets are publicly available and no agreed quality standards exist on the personalization of learning. The SIG dataTEL aims to address these issues by advancing data driven research to gain verifiable and valid results and to develop a body of knowledge about the personalization of learning.

Owner: Hendrik Drachsler
Group members: 66

The growth of data in the knowledge society creates opportunities for new insights through advanced analysis methods based on information retrieval technologies. Educational institutions also create and own huge datasets on their students and course activities. But they make little use of the data when considering new educational services, recommending suitable peers or content, and improving the personalization of learning.

The SIG dataTEL aims to address these issues by advancing data-driven research to gain verifiable and valid results and to develop a body of knowledge about the personalization of learning. It builds upon the positive outcomes of the dataTEL Theme Team funded by the STELLAR Network of Excellence. It’s intentions are to foster the cooperation between different Learning Analytics research units and to act as their representative to other relevant communities.

Therefore, its main objectives are:

**Networking:**
- Fostering of a research network on educational dataset driven research
- Improving the exchange with relevant research communities
- Representing dataTEL researchers to promote the release of open datasets from educational providers

**Privacy and Ethics:**
- Contributing to policies on ethical implications (privacy and legal protection rights)
- Suggesting guidelines for the anonymisation of data and
dataTEL - EATEL SIG on Data-driven Research and Learning Analytics

The growth of data in the knowledge society creates opportunities for new insights through advanced analysis methods based on information retrieval technologies. Educational institutions also create and own huge datasets on their students and course activities. But they make little use of the data when considering new educational services, recommending suitable peers or content, and improving the personalization of learning.

The SIG dataTEL aims to address these issues by advancing data-driven research to gain verifiable and valid results and to develop a body of knowledge about the personalization of learning.

Owner: Hendrik Drachsler
Group members: 66

Networking:
- Fostering of a research network on educational dataset driven research
- Improving the exchange with relevant research communities
- Representing dataTEL researchers to promote the release of open datasets from educational providers

Privacy and Ethics:
- Contributing to policies on ethical implications (privacy and legal protection rights)
- Suggesting guidelines for the anonymisation of data and
- Ensuring the legal and ethical use of data
Objectives of the SIG d

• Fostering of a research network on educational datasets

• Representing dataTEL researchers to promote the release of open datasets

• Contributing to policies on ethical implications (privacy and legal protection rights)

• Fostering a shared understanding of evaluation methods in Learning Analytics

• Fostering the standardizations of datasets to enable exchange and interoperability
How to achieve the goals

• Setting up a website / maintain TELeurope group community

• Set up a open data repository for sharing educational datasets and related researches outcomes

• Organizing annual workshop and SI

• Organizing a data competition like in TREC
HOW TO JOIN THE SIG dataTEL?
You can join the SIG dataTEL by becoming a member of the group space at the TELeurope.eu community platform.
Over the dataTEL group space we disseminate our activities, prepare workshops, share information, and discuss pressing issues.
If you care about educational data and it’s related research topics join us on TELeurope: http://bit.ly/datatel

SIG dataTEL Chairs:
Hendrik Drachsler, Centre for Learning Sciences and Technologies (CELSTEC)
Open University of the Netherlands
Tel. 0031-(0)45-57 62 218
Email: hendrik.drachsler@ou.nl

Katrien Verbert, Dept. Computer Science,
K.U.Leuven, Belgium
Tel: 0032-(0)16-327 060
Email: katrien.verbert@cs.kuleuven.be

CALL FOR JOURNAL PAPERS
Special Issue on dataTEL
“Datasets and Data Supported Learning in Technology-Enhanced Learning”
International Journal of Technology Enhanced Learning (IJTEL)
ISSN (Online): 1753-5263
ISSN (Print): 1753-5255

Deadline of submissions: 25 October 2011

TOPICS
Relevant topics include, but are not limited to:
- descriptions of datasets that can be used for experimentation
- descriptions of data experiments (methods or results of experiments)
- experiences with those datasets
- dealing with legal protection rights towards datasets on a European level
- privacy preservation for educational datasets
- methods of effective anonymisation of educational datasets
- management and pre-processing procedures for educational datasets
- future scenarios for educational datasets
- impact of educational datasets for learners, teachers, and parents
- mash-ups based on educational datasets
- recommender approaches that are based on educational data
- evaluation methodologies and metrics for educational recommender systems

SPECIAL ISSUE CO-EDITORS
Hendrik Drachsler, Open University, Netherlands
Katrien Verbert, K.U. Leuven, Belgium
Miguel-Angel Sicilia, University of Alcalá, Spain
Nikos Manouselis, Agro-Know Technologies, Greece
Stefanie Lindstaedt, KnowCenter, Austria
Martin Wolpers, Fraunhofer Institute for Applied Information Technology, Germany
Riina Vuorikari, European Schoolnet, Belgium

SUBMISSIONS
Submission will be available through the EasyChair submission system:
http://www.easychair.org/conferences/?conf=datatel2011

IMPORTANT DATES
Submission of manuscripts: 25.10.2011
Completion of first review: 30.11.2011
Submission of revised manuscripts: 15.01.2012
Final decision notification: 10.02.2012
Publication date (tentative): February 2012
Call for Journal Papers (IJTEL)

Datasets and Data Supported Learning
Deadline: 25.10.2011


SIG dataTEL Chairs:
Hendrik Drachsler, Centre for Learning
Sciences and Technologies (CELSTEC)
Open University of the Netherlands
Tel: 0031-(0)45-57 62 218
Email: hendrik.drachsler@ou.nl

Katrien Verbert, Dept. Computer Science,
K.U.Leuven, Belgium
Tel: 0032-(0)16-327 060
Email: katrien.verbert@cs.kuleuven.be

- privacy preservation for educational datasets
- methods of effective anonymisation of educational datasets
- management and pre-processing procedures for educational datasets
- future scenarios for educational datasets
- impact of educational datasets for learners, teachers, and parents
- mash-ups based on educational datasets
- recommender approaches that are based on educational data
- evaluation methodologies and metrics for educational recommender systems

EUROPEAN ASSOCIATION OF EATEL
TECHNOLOGY ENHANCED LEARNING

FRIDAY 02 SEPTEMBER 2011
TO SATURDAY 03 SEPTEMBER 2011
Publication date (tentative): February 2012
Learning Analytics Questionnaire

http://bit.ly/Learning_Analytics

- 234 response
- 16 countries
- Mostly Higher Education
Many thanks for your attention, and now...

picture by Tom Raftery  http://www.flickr.com/photos/traftery/4773457853/sizes/l
Many thanks for your attention, and now...

You face ethical and privacy issues when you want to work with edu. data!

picture by Tom Raftery  http://www.flickr.com/photos/traftery/4773457853/sizes/l
Many thanks for your attention, and now...

You are looking for the right tools to manipulate your data!

You face ethical and privacy issues when you want to work with edu. data!

picture by Tom Raftery  http://www.flickr.com/photos/traftery/4773457853/sizes/l
Many thanks for your attention, and now...

You are looking for educational datasets!

You are looking for the right tools to manipulate your data!

You face ethical and privacy issues when you want to work with edu. data!

picture by Tom Raftery  http://www.flickr.com/photos/traftery/4773457853/sizes/l
Many thanks for your attention, and now...

You are looking for educational datasets!

You face ethical and privacy issues when you want to work with edu. data!

You are looking for the right tools to manipulated your data!

picture by Tom Raftery  http://www.flickr.com/photos/traftery/4773457853/sizes/l