

# dataTEL - Final Report

Citation for published version (APA):

Drachslar, H. (2011). *dataTEL - Final Report*.

## Document status and date:

Published: 07/12/2011

## Document Version:

Peer reviewed version

## Document license:

CC BY-NC-SA

## 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](mailto:pure-support@ou.nl)

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 04 Dec. 2024

Open Universiteit  
[www.ou.nl](http://www.ou.nl)





## FINAL REPORT

### STELLAR Theme Team Programme

Deadline: **31<sup>st</sup> October 2011**

Dear Theme Team Leader,  
in order to help STELLAR to evaluate the Theme Team Programme as an integration instrument, you are kindly requested to provide a report of the scientific activities and outcomes you have carried out, starting from what was described in your proposal.  
Please send the document back to ITD-CNR (Francesca Pozzi [pozzi@itd.cnr.it](mailto:pozzi@itd.cnr.it)) by **October, 31<sup>st</sup>, 2011**. Thank you for your collaboration!

Your name: Hendrik Drachsler

Your Theme Team: dataTEL

Please give account of any change occurred in the Theme Team partnership: None

#### **GENERAL INFORMATION** (2 pages maximum).

Starting from what was indicated in your proposal, please describe your experience.

Be sure to include ALL the activities and events (even if you have already accounted some of them in the intermediate report).

#### **1. Activities**

- Created a living dataTEL community at TELEurope.eu (<http://www.teleurope.eu/pg/groups/all/?filter=pop>)
- Various new connections to relevant research communities (Movielens, Mendeley, Educational data mining, PSCL datashop, CAM and KDD data competition, CEN PT Social Data)
- Bi-weekly FM meetings with interested people from the field and related EU projects
- Organized three workshops with four keynote speakers from related research communities:
  1. RecSysTEL (ECTEL11, ACM RecSys11)
  2. dataTEL (ARV2011)
  3. dataTEL and LinkedEducation (LAK12 – proposal submitted)
- Organized a dataTEL Cafe' at EC-TEL 2010, and a dataTEL round table at ICL 2010
- Motivated Mendeley to release a first dataset for research (<http://dev.mendeley.com/datachallenge/>)
- Created an initial collection of TEL datasets (see article Verbert et al. submitted)
- Recorded four podcasts for the STELLAR podcasts series
- Initiated a Special Issue series on dataTEL at IJTEL Journal
- Conducted a Learning Analytics stakeholder questionnaire

#### **2. Main results**

##### **2.1 Publications**

1. Drachsler, H., Bogers, T., Vuorikari, R., Verbert, K., Duval, E., Manouselis, N., Beham, G., Lindstaedt, S., Stern, H., Friedrich, M., & Wolpers, M. (2010). Issues and Considerations regarding Sharable Data Sets for Recommender Systems in Technology Enhanced Learning. Elsevier Procedia Computer science 2010, pp. 2849-2858.

2. Verbert, K., Drachsler, H., Manouselis, N., Wolpers, M., Vuorikari, R., & Duval, E. (2011). Dataset-driven Research for Improving Recommender Systems for Learning. 1st International Conference Learning Analytics & Knowledge. February, 27 - March, 1, 2011, Banff, Alberta, Canada.
3. Drachsler, H., Verbert, K., Manouselis, N., Wolpers, M., Vuorikari, R., Lindstaedt, S. (2011). dataTEL – Datasets for Technology Enhanced Learning. STELLAR Alpine Rendez-Vous White Paper. Accessible at:  
<http://oa.stellarnet.eu/openarchive/browse?browse=collection%2Fall%2Fpublication&filter=all&param=69>
4. Greller, W., & Drachsler, H. (submitted). Turning Learning into Numbers – A learning analytics Framework. International Journal of Educational Technology & Society.
5. Verbert, K., Manouselis, N., Drachsler, H., Duval, E. (submitted). Dataset-driven Research to Support Learning and Knowledge Analytics. International Journal of Educational Technology & Society.
6. Drachsler, H., & Greller, W. (submitted). Confidence in Learning Analytics. 2nd International Conference Learning Analytics & Knowledge. Vancouver, Canada.

## 2.2 Workshops / Face-to-Face meetings

Date	Location	Type <sup>1</sup>	Activities	Outcomes	Links / Docs
17.09.2010	Round table, ICL conference 2010, Hasselt, Belgium	W	20 minutes presentation of the dataTEL objectives and 45 minutes open discussion.	-Privacy issues are the major challenge esp. in industrial and university settings	<a href="http://www.icl-conference.org/">http://www.icl-conference.org/</a>
28.09. – 29.10.2010	RecSysTEL workshop, ECTEL and ACM RecSys conference 2010, Barcelona Spain	W	Knowledge exchange on most pressing topics	- Need for a standardized meta-data structure for datasets - Privacy and legal protection regulations - Anonymization and pre-processing guidelines - Shared evaluation metrics to compare findings	<a href="http://adenu.ia.uned.es/workshops/recsyste12010">http://adenu.ia.uned.es/workshops/recsyste12010</a> , <a href="http://www.drachsler.de/?p=441">http://www.drachsler.de/?p=441</a> , <a href="http://dev.mendeley.com/datchallenge/">http://dev.mendeley.com/datchallenge/</a>
29.10.2010	DataTEL Café at the ECTEL conference 2010, Barcelona,	M	Discussions about data supported education during Coffee breaks and dinner sessions	Awareness raising in the community	<a href="http://www.ectel2010.org/programme">http://www.ectel2010.org/programme</a>

<sup>1</sup> W=Workshop, M=Meeting, C=Conference

	Spain				
27-29.03.2011	DataTEL - Datasets for Technology Enhanced Learning at ARV2011, La Clusaz, France	W	Focused session on 4 pressing topics: 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	4 Grand Challenges for data supported education described in the dataTEL white paper.	<a href="http://www.tel-europe.eu/pg/pages/view/46082/">http://www.tel-europe.eu/pg/pages/view/46082/</a>  White paper: <a href="http://oa.stellarnet.eu/open-archive/browse?browse=collection%2Fall%2Fpublication&amp;filter=all&amp;param=69">http://oa.stellarnet.eu/open-archive/browse?browse=collection%2Fall%2Fpublication&amp;filter=all&amp;param=69</a>
21.09.2011	Launch of SIG dataTEL at ECTEL conference 2011, Palermo, Italy	C	Presentation of Special Interest Group "Data Driven Research and Learning Analytics" (DataTEL) of EATEL	Awareness raising in the community	<a href="http://www.ec-tel.eu/programme">http://www.ec-tel.eu/programme</a>
29.04.2012	dataTEL, Learning Analytics conference in Vancouver, Canada	W	Exploring the potential of Linked Data and Learning Analytics workshop	-	-

### 3. Face-to-Face meetings

The dataTEL Theme Team mainly had biweekly FM meetings and met at the workshops and other scientific events Face-to-Face. Next to the internal meetings some members of the Theme Team visit each other at their local research units.

### 4. Changes to original plans:

There have been several extensions and also limitations to the original research plan. Overall it is to mention that we contributed to all five main research topics (1. Privacy legal protection, 2. Data sharing, 3. Processing data, 4. Evaluation criteria for data products, 5. Overview methods to findings on datasets) described in the project proposal. During the runtime of the Theme Team we were able to define the underlying practical problems much clearer. Privacy and legal protection remain the biggest challenges for dataTEL research. But through the scientific events we were able to join forces with several related projects and data scientists to develop research plans for upcoming research proposals. Furthermore, the focus of the Theme Team has been broaden up from a pure recommender benchmarking data perspective to also reflection and awareness support by educational data

products that are studied in the field of Learning Analytics and other data supported research initiatives like Linked data, educational data mining, and Research2.0.

## 5. **Sustainability:**

In order to sustain the positive outcomes of the dataTEL Theme Team and to continue the work on the dataTEL challenges on a project independent and international level, we initiated the Special Interest Group dataTEL under the umbrella of the European Association of Technology Enhanced Learning (EATEL). It's intentions are to foster the cooperation between different Learning Analytics research units and to act as their representative to other relevant communities. The SIG dataTEL is intended as a network to collect, validate, and discuss different Learning Analytic approaches based on datasets from real educational settings. Therefore, its main objectives are:

- Fostering of a research network on educational dataset driven research and improve the exchange with relevant research communities
- Representing dataTEL researchers to promote the release of open datasets from educational providers
- 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
- Making dataTEL related research results from EU projects sustainable

To facilitate the SIG dataTEL, it is intended to continue the existing dataTEL group on TELeurope and to organize one dataTEL workshop per year at leading conferences in the field. As the next step, we propose to create a website for the SIG and to set up an open repository with publicly available datasets and dataTEL related research results.

## ***SCIENTIFIC ADVANCES*** (2 pages maximum).

In this section you are requested to describe the scientific advances made by your Theme Team, especially in the light of the STELLAR Grand Challenges framework (see the Call).

### **1. briefly describe the specific topic(s) tackled in your Theme Team**

In the dataTEL Theme Team we addressed the following topics:

- publicly available datasets for educational systems
- dealing with legal protection rights towards datasets on a European level
- privacy preservation for educational datasets
- procedures for effective anonymisation of educational datasets
- management and pre-processing procedures for educational datasets
- future scenarios for educational datasets
- impact of educational datasets for learners and teachers
- mashups based on educational data sets
- recommender approaches that are based on educational data
- evaluation methodologies and metrics for educational recommender systems

### **2. state CLEARLY the main research questions your Theme Team has identified**

The dataTEL Theme Team clearly investigated a couple of research questions that we clustered according to 4 most pressing topics:

#### 2.1 Evaluation of recommender systems in TEL

- Does a common data format for evaluation also require a common format for TEL recommender systems?
- Learning is a collaborative process, how can we translate that into evaluation measures for TEL recommender systems?
- Do we have to converge to a common evaluation framework or should we diverge to a wide range of (accepted) evaluation methods to choose from?
- We need metrics that can be applied to every standardized dataset!

## 2.2 Data supported learning examples

- How can we get richer representations of the social context in learning?
- How important are visualisations to reflect and learn from a dataset?
- Can teachers or students deal with data visualisations to reflect their learning process?
- Which new competences requires data supported learning?
- How can we integrate the context of learning into the support systems?

## 2.3 Datasets from Learning Object Repositories and Web content

- Based on the experiences with SCORM and IMS-LD that should have created a European Learning Object market, will a common dataset format really lead to more datasets in TEL?
- How can we create datasets that capture real-life learner data?
- How can we overcome the lack of data sharing opportunities?
- Do we need a dataset format or rather well documented datasets? or both?
- How can we deal with the diversity of data from various TEL systems? What are appropriate levels of granularity?
- We need to create a representative association that requests datasets from the big players in LMS (e.g., Moodle, Blackboard) and learning object repositories (e.g., MERLOT, OERCommons)

## 2.4 Privacy and data protection for dataTEL

- Does dataTEL require an ethical discussion on privacy, data protection and surveillance?
- We are able to develop a new generation of support tools when we are allowed to track the context and behaviour of learners. How can we deal with privacy issues in a practical way?
- Data driven research will make unveil information visible that will challenge the way we learn, teach and conducting research.

### 3. Addressing research questions by defining Grand Challenge Problems

All of the above mentioned questions are relevant and offer a future research agenda for the EATEL SIG dataTEL and several EU projects. At the dataTEL workshop at the ARV11 we also defined 4 most Grand Challenges out of the pressing topics, we describe below in more detail.

#### **Grand Challenge 1 / Topic 1:**

***Reduce the drop-out rate in online learning environments by 10% through applying well evaluated and tested recommender systems for learning.***

*What problems of the European education system are addressed, and what are the long term benefits for society?*

A challenging problem for educational institutes and lifelong learning in general are the high drop-out rates esp. in online and distance education settings. The isolation and confusion of students may cause them to withdraw from their studies. These groups of students are called 'drop-outs'. The research on TEL recommender systems can contribute to decrease the drop-out rate by disseminating its research outcomes for the development of different support systems for teachers and students to offer relevant information at the right time. Regarding the drop-out problem it is thinkable to develop a drop-out analyzer that informs the tutor of a (Moodle) course which learners are likely to drop-out. This could be done by training a certain recommender technology on the drop-out patterns of previous (Moodle) courses. The trained analyzer could then be applied on follow-up (Moodle) courses and mark students in a list that show similar drop-out patterns. The tutor of the course could then make an intervention and contact those students personally to offer additional support for their studies.

*What are the main activities to address this Grand Challenge Problem?*

- Customize existing recommendation algorithms for learning
- Employ recommender systems in real-life scenarios
- Develop suitable evaluation criteria for different kind of recommender systems

*What is the timeframe for the Grand Challenge Problem?*

First implantations recommender systems are already available and can be implemented within a year (Manouselis et al. 2011; Manouselis et al., 2010). More challenging is the evaluation of the recommender systems that will take up to 2 – 3 years. For the further development of such systems publicly available educational datasets are needed to evaluate and compare different recommendation approaches to gain a solid body of knowledge (5-8 years).

*What are measurable progress and success indicators?*

Measurable progress and success indicators are depending on the applied type of recommender system (curriculum recommender system, drop-out analyzer etc.) (Drachslar, Hummel, Koper, 2009). For the this Grand Challenge a significant decrease of the drop-out rate within an educational institution would be an promising measure to value the impact of such a system. A challenging issue will be to isolate the effect of decreasing drop-out rates only to the recommender system as most educational institutes permanently improve their educational services.

*How can funding be attracted?*

Next to European and national funding such a research project could be funded by single Universities (Innovation funds) and LMS providers like Blackboard or IMC AG. Next to the commercial providers recommender projects can be initiated as open source project in the Moodle or SAKAI community for instance.

### **Grand Challenge 2 / Topic 2:**

***ACTUALLY, help students and teachers in TEL to use data supported information systems.***

*What problems of the European education system are addressed, and what are the long term benefits for society?*

In order to make data supported information systems an effective tool for educational practice, various limitations and hurdles in technology, privacy and education need to be addressed. It is

important to realise that data supported tools work with computational results that are not easy to understand and need to be presented in an easy way (e.g., by visualizations) to address the daily practice of the educational stakeholders. It is crucial to interpret the presented outcomes in a correct manner to take the right follow-up activities that can lead to improved learning. Therefore, the interpretation of educational data and its related tools requires new competences to deal with the outcomes (statistical knowledge, critical thinking, privacy awareness and ethical competences).

*What are the main activities to address this Grand Challenge Problem?*

- Developing new data driven tools that are easy to understand
- Make new real time data tools available as test applications
- Identify suitable algorithms and map them to certain datasets and learning purposes
- Integrate statistical, critical thinking, privacy awareness and ethical competences into the teacher education programs

*What is the timeframe for the Grand Challenge Problem?*

The development of the data tools has already started, on every conference new data driven tools are presented (Zhang & Almeroth, 2010). Systems like Mendley and open access journals show us already the future of academic work. The training of the new competences for teachers and students in the primary and secondary education level will take more time (5 to 10 years) with having many different levels in the EU partner countries.

*What are measurable progress and success indicators?*

- An increased effectiveness, efficiency or satisfaction of the learning process
- Courses at educational providers that train competences to handle data products
- An increasing amount of data mashup systems for different educational stakeholders (students, teachers, parents and educational providers)

*How can funding be attracted?*

For the competence training the Lifelong Learning Programme of the EU is suitable. The development of the new data driven tools can be funded by FP7 and national calls.

### **Grand Challenge 3 / Topic 3:**

***Create a generic infrastructure for sharing, analyzing and reusing learning resources and learning activity logs (educational datasets) and related research findings.***

*What problems of the European education system are addressed, and what are the long term benefits for society?*

The increased application of LMS, e-portfolio systems, and PLEs in schools and higher education institutions produces large amounts of educational data. But, although these e-learning environments store educational data automatically, exploitation of this data for learning and teaching is still very limited. These educational datasets offer an unused opportunity for the evaluation of learning theories, student support, learning technology, and the development of future learning applications. Furthermore, educational datasets can be supportive to advance research on TEL towards a basic theory for TEL (Verbert et al. 2011) by offering the recorded



and observed behavior of the stakeholders (students, teachers, parents, lifelong learners, educational institutes) in different learning settings (formal – informal learning). In that way, the educational datasets extend the methodological and empirical approaches to analyze TEL that is dominated by design-based research approaches, simulations, and field studies (Gray, 2009).

*What are the main activities to address this Grand Challenge Problem?*

- Data ownership and access rights are challenging because the LMS and PLE systems are collecting educational data and the current assumption is that this data belongs to them. However, who exactly holds the ownership of the data created by the students and what can be done with it is still unresolved.
- Data policies (licences) that regulate how different users can use, share, and reference certain datasets. Until now there are very limited data policies available in educational institutes. It could be considered to apply the Creative Commons licensing rights as a standard way to grant permissions to datasets.
- Common dataset formats like from the CEN PT Social data group (<https://sites.google.com/site/camschema/home>) and a standardised documentation of datasets so that others can make proper use of it.
- Methods to anonymise and pre-process data according to privacy and legal protection rights.

*What is the timeframe for the Grand Challenge Problem?*

Anything between 5 and to 8 years. For learning resources there are already standards like LOM and Dublin Core. For learning activities it's more complicated (apart from very generic formats such as XML – which does not guarantee that data can be reused).

*What are measurable progress and success indicators?*

- An increasing amount of publicly available datasets and research articles that are based on shared datasets
- The availability of data or privacy policies at educational providers
- More data-driven tools at educational providers
- A common dataset format

*How can funding be attracted?*

Funding can be attracted from governmental funding bodies like FP7, national funding, or funding by companies like Microsoft, Google, or IBM.

#### **Grand Challenge 4 / Topic 4:**

***Reduce delivery costs and create more effective learning environments by applying advanced information retrieval technologies on educational data sets.***

Europe's education systems suffer from decreasing amount of teachers and the request to increase the amount of high-educated students in a short time period. As a consequence there is less time available for the individual support of students, thus the teaching quality decreases. On the other hand, the education systems are increasingly based on electronic systems like LMS and e-portfolios. With the increase in available educational data, the application of information

retrieval technologies becomes valuable to create new services for the educational stakeholders (students, parents, teachers, and educational institutes). The combination of educational data and information retrieval techniques also known as Learning Analytics (LA) will become a powerful means in educational practice and student guidance (Johnson et al., 2011). LA promises the educational field to reduce delivery costs, create more effective learning environments and experiences, accelerate competence development, and increase collaboration between students and teachers.

But LA also have barriers and limitations among these are issues of privacy and data protection that need to be addressed by policy guidelines. Additional, challenges arise with respect to data surveillance<sup>2</sup> (social sorting, cumulative disadvantages) and its ethical implications.

*What are the main activities to address this Grand Challenge Problem?*

- In order to discuss and improve the above-mentioned situation a new vocabulary needs to be accomplished in order to discuss privacy, data protection and surveillance issues. For instance, what are better terms to express concepts like ownership and access control, when in digital systems replication and distribution is so easy that the concepts have no traction.
- Research is needed on how existing privacy and transparency solutions can be integrated in dataTEL practice. Further, research is desirable on how state of the art security solutions can be used to secure large educational datasets.
- There is a need for data awareness education for society. Such an educational program should not be limited to teaching individuals when to reveal or conceal their data, but also to increase their awareness with respect to large datasets, surveillance practices, and related problems.
- User and stakeholder studies (case studies) are necessary to understand the complex requirements with respect to privacy, data protection, surveillance in dataTEL.
- The issues around privacy, data protection, and surveillance need to be addressed from the beginning of the research and not as an add-on. Methodologies and guidelines that support this vision need to be developed to support privacy and ethical practices.
- There needs to be research on how to bridge between dataTEL researchers and ethical boards with respect to advances in technologies and research and the related privacy, data protection, and surveillance concerns that arise with them.
- Policies have to be defined to avoid unethical data mining research.

*What is the timeframe for the Grand Challenge Problem?*

---

<sup>2</sup> Data surveillance refers to the process which individualizes each member of the population (or a group), and permits the observation and recording of each individual's activities, then collates these individual observations across the population. From these conglomerated observations, statistical norms are produced relating to any of a multitude of characteristics. These norms are then applied back to the subjected individuals, who are categorized and perhaps acted upon, either with gratification or punishment, according to their relation to the produced norm. (Phillips, Privacy Policy and PETs, 2004)

The first four activities can be addressed in a time frame from 2 to 3 years because they mainly require the application or translation of existing examples or solution from other domains to the educational field. The activities 5 to 7 will require a longer timeframe (3-5 years) as they can only be developed out of the experiences with the activities 1 to 4.

*What are measurable progress and success indicators?*

Measureable progress and success indicators are an increasing amount of ethical boards in LA units at educational organisations. The integration of privacy and data protection statements in research projects as well as between educational providers and the students. The integration of data and privacy competence in job profiles at the educational providers.

*How can funding be attracted?*

Funding for these challenges could be attracted from EU FP7 projects and the Lifelong Learning Programmes (Erasmus, Leonardo or Comenius).

### ***EVALUATION (1 page maximum).***

Please provide your own evaluation of the overall Theme Team experience, especially in terms of:

*a) quality of interactions occurred within the Theme Team and contribution of individual partners*

The overall cooperation within the Theme Team was very positive, we had very fruitful discussions and inspiring outcomes. Although, the Theme Team funding did not provide staff costs the consortium could integrate the objectives of the Theme Team into their running and related research activities.

*b) level of scientific integration (e.g. co-authored publications, jointly organized events, memorandums of understanding between institutions)*

The positive cooperation in the Theme Team also affected the outcomes of the project. dataTEL went far beyond the expected outcomes described in the project proposal.

*c) quality of the outputs produced*

The quality of the output was very high in terms of innovation and visibility. We were able to attract several related research communities, improved the visibility of TEL research, made dataTEL a label for data driven research in TEL, received very good responses from the participants in the workshops and conducted one of the first dataset challenge in the TEL community.

*d) possible plans for the future*

There are several research plans for the close future. The outcomes and ideas of dataTEL will influence new EU projects and scientific events. With the SIG dataTEL at EATEL (see section sustainability) we have a clear roadmap for the upcoming two years for data supported research in TEL.

*e) interaction with the STELLAR partners (as far as the scientific aspects and the financial ones)*

The interaction with STELLAR partners was very fruitful, the clear focus on developing new Grand Challenges nicely harmonised with the most pressing topics defined in the dataTEL workshops. The ARV11 was a very good ground to further develop these ideas and also workout clear objectives for the future projects.

*f) any other comments (weaknesses or strong points) you want to provide regarding the Theme Team instrument.*

The Theme Teams are a very effective instrument to get fast and unconventional ideas disseminated and evaluated in the community. It was very supportive to have the opportunity to invite relevant experts from all over the world and fund scientific events without many administrative hurdles.

The only short coming is that the consortium need to have the opportunity to link a Theme Team activities to related research efforts otherwise the missing of staff budget could lead to decreasing participation.

### ***EXECUTIVE SUMMARY (half page maximum).***

Please provide a synthetic overview of the work done during your Theme Team and of the main result(s). Note that this summary will be made public through the STELLAR website.

The work of the dataTEL Theme Team has resulted in several important outcomes. An initial format for the collection of TEL datasets has been reported in (Drachslers et al. 2010). This format was used to collect an initial collection of datasets that capture learner interactions with tools and resources in real-life settings. These datasets were collected by the first dataTEL Challenge, launched as part of the workshop on Recommender Systems for TEL, jointly organized 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 datasets from TEL applications. Several datasets have been collected and first evaluation experiments with these datasets have been reported in (Verbert et al. 2011). As a follow up activity, the “dataTEL - Datasets for Technology Enhanced Learning” workshop was organized at the Second STELLAR Alpine Rendez-Vous in March 2011. During this workshop, researchers discussed related initiatives that are collecting educational datasets, additional datasets that are relevant for LAK research, as well as challenges related to privacy and data protection and research on evaluation methodologies. An analysis of related initiatives, dataset formats and available datasets has been reported in (Verbert et al. 2011a). The outcomes of the dataTEL workshop are going to be combined in a Special Issue at the IJTEL journal that will be publicly available in 2012. In order to continue the work on dataTEL research topics the SIG dataTEL was created under the umbrella of the European Association of Technology-Enhanced Learning (EATEL). The SIG dataTEL will continue the existing dataTEL group on TELEurope and organize an annual dataTEL workshop at leading conferences in the field.

### **References:**

- Drachslers, H., Bogers, T., Vuorikari, R., Verbert, K., Duval, E., Manouselis, N., Beham, G., Lindstaedt, S., Stern, H., Friedrich, M., & Wolpers, M. (2010). Issues and Considerations regarding Sharable Data Sets for Recommender Systems in Technology Enhanced Learning. Elsevier Procedia Computer science 2010, pp. 2849-2858.
- Drachslers, H., Hummel, H. G. K., & Koper, R. (2009). Identifying the Goal, User model and Conditions of Recommender Systems for Formal and Informal Learning. Journal of Digital Information, 10(2), 4-24.

- Drachsler, H., & Greller, W. (submitted). Confidence in Learning Analytics. 2nd International Conference Learning Analytics & Knowledge. Vancouver, Canada.
- Gray, J., (2009). A Transformed Scientific Method, In: The 4th Paradigm: Data Intensive Scientific Discovery.
- Greller, W., & Drachsler, H. (submitted). Turning Learning into Numbers – A learning analytics Framework. International Journal of Educational Technology & Society.
- Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H. and Koper, R. (2010). Recommender Systems in Technology Enhanced Learning. In Kantor, P.B.; Ricci, F.; Rokach, L.; Shapira, B. (Eds.) 1st Recommender Systems Handbook, Berlin: Springer (2011).
- Manouselis, N., Drachsler, H., Verbert, K., & Santos, C. S. (Eds.) (2010). Recommender System in Technology Enhanced Learning. Elsevier Procedia Computer Science: Volume 1, Issue 2. Proceedings of the 1st Workshop on Recommender Systems for Technology Enhanced Learning (RecSysTEL). September, 29-30, 2010, Barcelona, Spain. doi: 10.1016/S1877-0509(10)00329-7.
- Verbert, K., Drachsler, H., Manouselis, N., Wolpers, M., Vuorikari, R., & Duval, E. (2011). Dataset-driven Research for Improving Recommender Systems for Learning. 1st International Conference Learning Analytics & Knowledge. February, 27 - March, 1, 2011, Banff, Alberta, Canada.
- Verbert, K., Manouselis, N., Drachsler, H., Duval, E. (2011a). Dataset-driven Research to Support Learning and Knowledge Analytics. International Journal of Educational Technology & Society (submitted).
- Johnson, L., Smith, R. Willis, H. Levine, A. Haywood, K. (2011). The 2011 Horizon Report. Austin, Texas: The New Media Consortium.
- Zhang, H. and Almeroth, K. (2010). Moodog: Tracking Student Activity in Online Course Management Systems. Journal of Interactive Learning Research, 21(3), 407-429. Chesapeake, VA: AACE. Retrieved May 1, 2011 from <http://0-www.editlib.org.aupac.lib.athabascau.ca/p/32307>.