Leveraging Learning Analytics in a Personal Learning Environment using Linked Data

Selver Softic, Laurens De Vocht, Behnam Taraghi, Martin Ebner, Erik Mannens and Rik V. De Walle

Abstract—**We report on the reflection of learning activities and revealing hidden information based on tracking user behaviors with Linked Data. Within this work we introduce a case study on usage of semantic context modelling and creation of Linked Data from logs in educational systems like a Personal Learning Environment (PLE) with focus on reflection and prediction of trends in such systems. The case study demonstrates the application of semantic modelling of the activity context, from data collected for over two years from our own developed widget based PLE at Graz University of Technology. We model learning activities using adequate domain ontologies, and query them using semantic technologies as input for visualization which serves as reflection and prediction medium as well for potential technical and functional improvements like widget recommendations. As it will be shown, this approach offers easy interfacing and extensibility on technological level and fast insight on trends in e-learning systems like PLE.**

Index Terms—**Data Mining, Semantic Web, Electronic learning, Analytic models**

I. Introduction

Limited availability of resources along with a time efficiency focus forces the designers and decision makers of learning platforms to revise their methodologies and techniques in order to respond the challenges of time and the needs of their targeted groups. On the other hand, learners are expecting a focused and simple way to organize their learning process, without losing time on information and actions which could disturb or prolong their learning. Nowadays learning process became more individual, multi-faceted and activity driven with the tendency to ad hoc initiated collaboration and information exchange. These circumstances imply the need for a scalable, adaptive learning environment enriched with multimedia supportive materials, communication channels, personalized search and interfaces to external platforms from Social Web like e.g. Slideshare, Youtube channels etc. All these parameters increase the complexity of online learning platform design and organization. Dynamics involved in this process require shorter optimization cycles in adaptation of learning process. Also maintaining such platforms is intensively changing process demanding from maintainers to actively adapt their systems to the learner needs. Adaptation to learner needs has a strong impact on acceptance of such platforms and should be matter of continuous improvement.

Cumulated system monitoring data (e.g. logs) of such environments offer new opportunities for optimization [1]. Such data can contribute the better personalization and adaptation of the learning process but also improve the design of learning interfaces.

Main contribution of the paper is a case study done with the logs from PLE at Graz University of Technology, presenting approach using Linked Data to mine the usage trends from PLE. The idea behind this effort is aiming at gaining insights, [2] useful for optimization of PLE and adapting them to the learners by using more personalization e.g. through recommendation of interesting learning widgets.

The research activities that have been described in this paper were funded by Graz University of Technology, Ghent University, iMinds (an independent research institute founded by the Flemish government to stimulate ICT innovation), the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT), the Fund for Scientific Research-Flanders (FWO-Flanders), and the European Union.

S. Softic, B. Taraghi and M.Ebner is with Social Learning Group, Graz University of Technology, Münzgrabenstraße 35A, 8010 Graz, Austria (corresponding author e-mail: selver.softic@ tugraz.at).

L. De Vocht, E. Mannens and R. Van De Walle is with iMinds - Multimedia Lab, Ghent University, Gaston Crommelaan 8, 900 Ghent Belgium (e-mail: first.last @ ugent.be).

II. Related Work

This section reports shortly about most relevant related work regarding PLE (at Graz University of Technology) and semantic technologies used in this work.

A. Learning Analytics and Reflection of Learner Logs

Current Learning Analytics research community defines [3] Learning Analytics as the analysis of communication logs [4], [5], learning resources [6], learning management system logs as well existing learning designs [7],[8] and the activity outside of the learning management systems [9],[10]. The result of this analysis improves the creation of predictive models [11], recommendations [12],[13] and refection [14]. Learning Analytics resides on algorithms, formulas, methods, and concepts that translate data into meaningful information. Modelling, structuring and processing the collected data derived from e.g. learner behavior tracking plays a decisive role for the evaluation. Different works outlined the importance of tracking activity data in Learning Management Systems [2],[3],[4],[12],[14]. None of them addressed the issue of intelligently structuring learner data in context and processing it to provide a flexible interface that ensures maximum benefit from collected information.

B. PLE at Graz University of Technology

The main idea of PLE at Graz University of Technology (http://ple.tugraz.at) is to integrate existing university services and resources with services and resources from the World Wide Web in one platform and in a personalized way [15], [16]. The TU Graz PLE contains widgets [15-17] that represent the resources and services integrated from the World Wide Web. Web today provides lots of different services; each can be used as supplement for teaching and learning. The PLE has been redesigned in 2011, using metaphors such as apps and spaces for a better learner-centered application and higher attractiveness [18],[19]. In order to enhance PLE in general and improve the usability as well as usefulness of each individual widget a tracking module was implemented by prior work [20].

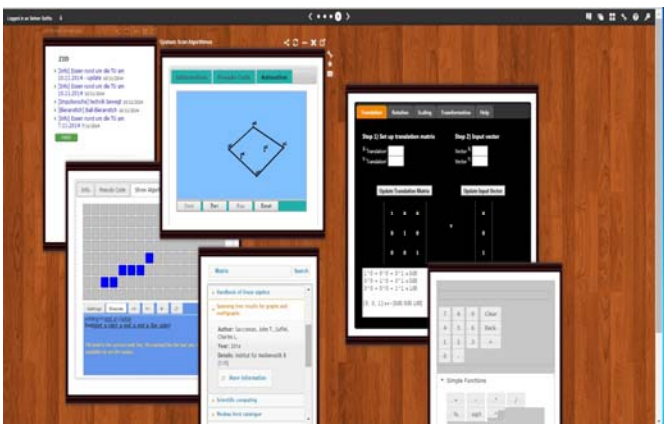

Fig. 1. PLE at Graz University of Technology.

C. Semantics for Modeling Learners in PLE

The Semantic Web standards like RDF (http://www.w3.org/RDF), RDFS (http://www.w3.org/TR/rdf-schema/), OWL (http://www.w3.org/2004/OWL/) and SPARQL (http://www.w3.org/TR/rdf-sparql-query/) enable data to be modeled and queried as graphs. Data schema is usually projected on specific knowledge domain using adequate ontologies. This approach has been fairly successful used to generate correct interpretation of web tables [21] to advance the learning process [1],[22] as well to support the controlled knowledge generation in E-learning environments [23]. This potential was also recognised by resent research in IntelLEO Project (http://intelleo.eu). IntelLEO delivered an ontology framework where Activities Ontology (http://www.intelleo.eu/ontologies/activities/spec/) is used to model learning activities and events related to them. The second interesting contribution from recent Ontology research work in IntellLEO project is the Learning Context Ontology which describes the context of a learning situation (http://www.intelleo.eu/ontologies/learning-context/spec/). Due to the relatedness to the problem that is addressed by this work those ontologies have been used to model the context of analytic data collected from PLE logs.

III. Approach for Mining Learner Logs

A. Dataset

Data used in the case study originates from Personal Learning Environment (PLE) developed for the needs of Graz University of Technology which serves currently more than 4000 users. The data was collected during two years period in order to generate analytics reports with visualization support for overall usage and process view on our environment following the research trends of previous years [3],[9].

B. Modeling Learner Logs

The main precondition for meaningful mining of usage trends is meaningful modelling of data. Such action assumes the choice of appropriate vocabulary or ontology. The RDF standard as such offers only the generic framework how to: organize, structure and link data.

The Activities Ontology introduced by IntelLEO research project offers a vocabulary to represent different types of activities and events related to them. Further this ontology also supports the description of environment (in our case PLE) where these activities occur. The Learning Context ontology serves as shown in figure 2. as container model to link PLE usages as event to the widget as execution environment where this event happens.

Every usage of a PLE widget creates a logging entry which produces a RDF construct represented in figure 3. Such constructs represent Linked Data which is then stored in a RDF memory store: Graph Database for Linked Data with SPARQL Endpoint, an interface where Linked Data can be queried.

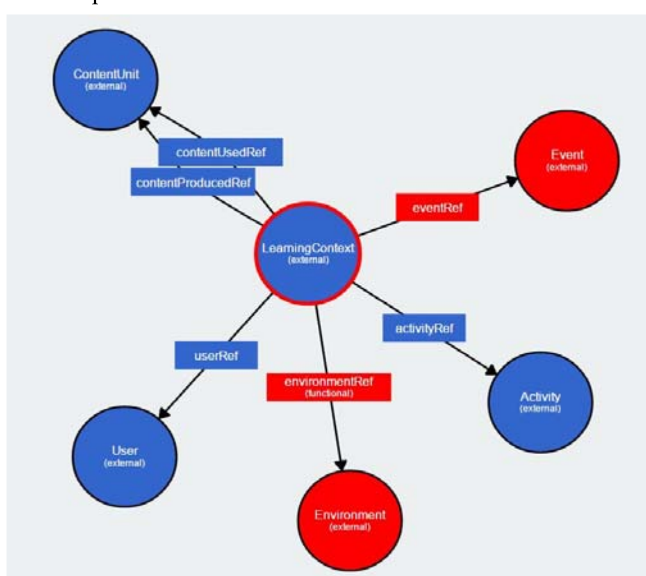

Fig. 2. Visualisation by WebVOWL beta 3.0 of a LearningContext ontology concepts and properties used to model the PLE log data.

```
@prefix ao: <http://intelleo.eu/ontologies/activities/ns/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix lc: <http://www.intelleo.eu/ontologies/learning-context/ns/> .

<http://ple.tugraz.at/ns/events/log/#7912>
    rdf:type ao:Logging;
    ao: occursIn <https://ple.tugraz.at/ns/widgets/#LatexFormulaToPngWidget>;
    ao:timestamp "2012-10-04T07:52:52" .

<https://ple.tugraz.at/ns/widgets/#LatexFormulaToPngWidget>
    rdf:type ao:Enivironment;
    rdfs:label "LaTeXFormulaPNG Converter" .

<http://ple.tugraz.at/ns/learningcontext/#7912> a lc:LearningContext;
    lc:environmentRef
    <https://ple.tugraz.at/ns/widgets/#LatexFormulaToPngWidget>;
    lc:eventRef <http://ple.tugraz.at/ns/events/log/#7912>.
```

Fig. 3. Sample model of a PLE log in expressed as LearningContext.

In figure 3. a sample instance of lc:LearningContext links the usage log event denoted as instance of class ao:Logging , a subclass of an ao:Event, which occurred at certain time point inside the learning widget named LatexFormulaToPNGWidget represented through class ao:Enviroment.

What we can see with this example is that: vocabularies and ontologies which suits well to specific use case, enrich the analytic process with a high level of expressiveness in a very compact manner.

C. Querying the Models

Usage logs data presented as Linked Data graph are query able using SPARQL. In this way we are able to answer the questions like: *"Show me the top 10 used widgets?"*. Figure 4. represents exactly this question stated in the manner of SPARQL syntax. The benefit of this query is visible for instance in figure 6. where the results of this query (see figure 5.) influence the widget arrangement in the widget store. Such direct impact on system with functional operability on machine level would not be possible without standards like SPARQL and RDF.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
PREFIX lc: <http://www.intelleo.eu/ontologies/learning-context/ns/>

SELECT DISTINCT ?widget ?ref (COUNT(?ref) AS ?count)
WHERE
{
 ?link rdf:type lc:LearningContext ;
       lc:environmentRef ?ref .

 ?ref rdfs:label ?widget .
}
GROUP BY ?ref
ORDER BY DESC(?count)
LIMIT 10
```

Fig. 4. Querying the usage of top 10 widgets in PLE.

IV. Conclusion and Outlook

Presented approach allows us mining the trends of PLE widgets usage overall time periods like presented in figure 5. This pie chart graph depicts the visual answer of the query from figure 4. The overview over distribution of widget usage can reflect the overall interest of the users within PLE for different periods of time. Such outputs implicitly support the improvement of the quality of services for students and teachers. The same results from query in figure 4. are also used as input for ranking of widgets in widget store depicted in figure 6. This example shows the manifold application of such approach. The PLE becomes, in technical manner, extensible and well connected by standardized and intelligent interfaces and available for other web based tools and services.

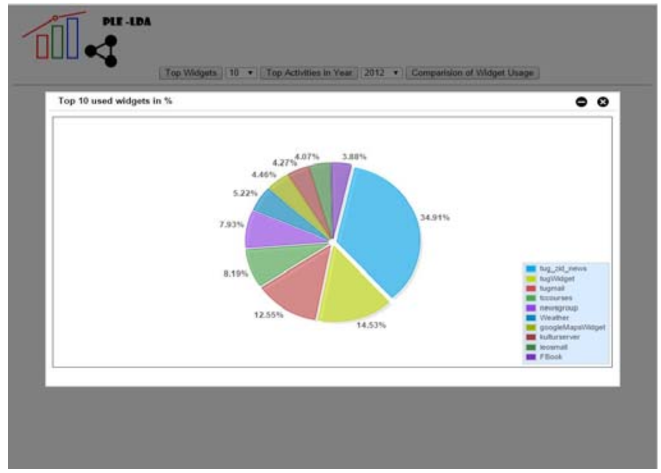

Fig. 5. Visualising top 10 used widgets in PLE.

Future efforts will focus on user wise statistics of learning widgets, since PLE can also provide this information. Especially the learning widget store as part of PLE could profit from this improvement. Mostly used and favored widgets by users will be ranked higher and recommended by the store itself as shown in figure 6. This process will be personalized as soon as the user information is included. The presented Learning Context ontology as such have foreseen such option already. By tracking the usages on user level the teachers will be able to draw conclusions about the popularity and quality of their learning widgets, on more granular and personal level.

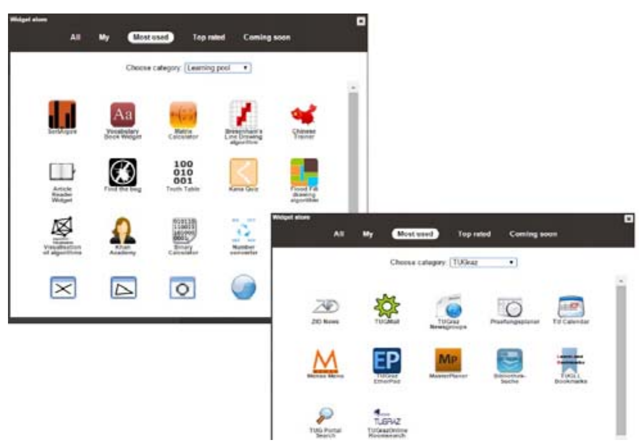

Fig. 6. Optimized widget store based on Linked Data statistics.

Beside presented practical benefits from using Linked Data in Learning Analytics demonstrated on the example of PLE it is important to pinpoint how such approach differs from conventional methods and to which extent extends them. While conventional Learning Analytics focus more on Monitoring and drawing conclusions from analyzed and derived results Linked Data driven Learning Analytics approach delivers and derives the results on demand and in in-time. This option offers technical extensibility and claims interoperability by default, opening the interfaces toward other web platforms, sources and internet technologies. Flexibility through SPARQL standard for interactive querying allows the dynamic generation of inputs for hosting platforms, visualizations and analytic dashboards.

Such actions require data models with certain degree of expressiveness, and well-thought-out constrains. Main challenge lies in choice or construction of proper model, as well as in the decision about the granularity degree of chosen model. Sometimes, this process is limited by the quality and variety of provided data. Very important advantage of such models is their adaptability to extensions, reductions and changes of model schema. The nature of RDF, RDFS and OWL allows also to inference based on logical rules. This is especially useful for asking sophisticated questions about the context of modeled data.

Leveraging Learning Analytics with Linked Data we support standardized interfaces for information exchange, offer flexibility for visual other kinds of analytics, and also can enrich the learning system's data with Linked Data sources from the Web as well with results from querying the graphs. The spread of applicability covers wide range of analytic methodologies like prediction, reflection and as outcome of these the intervention field.

Presented work based on case study of data from a PLE outlines the contribution of Semantic Technology Stack and Linked Data to Learning Analytics. The idea is promising and delivers great results with very low effort, what makes it especially valuable for analytical tasks targeted on improvement of learning environments like PLE.

REFERENCES

[1] P. Prinsloo, S. Slade and F. Galpin, "Learning analytics: challenges, paradoxes and opportunities for mega open distance learning institutions", Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, ACM, 2012, pp. 130-13.

[2] R. Mazza and C. Milani, "Exploring Usage Analysis in Learning Systems: Gaining Insights From Visualisations Workshop on Usage Analysis in Learning Systems", Proceedings of Artificial Intelligence in Education, 2005.

[3] J. L. Santos Odriozola, K. Verbert, S. Govaerts and E. Duval, "Visualizing PLE usage", Proceedings of EFEPLE11 1st Workshop on Exploring the Fitness and Evolvability of Personal Learning Environments, CEUR WS, 2011, pp. 34-38.

[4] D. Rosen, V. Miagkikh and D. Suthers, "Social and semantic network analysis of chat logs", Proceedings of the 1st International Conference on Learning Analytics and Knowledge, ACM, 2011, pp. 134-139.

[5] A. Bakharia and S .Dawson, "SNAPP: a bird's-eye view of ter participant interaction", Proceedings of the 1st International Conference on Learning Analytics and Knowledge, ACM, 2011, pp.168-173.

[6] K. Niemann, H. C. Schmitz, M. Scheffel and M. Wolpers, "Usage contexts for object similarity: exploratory investigations", Proceedings of the 1st International Conference on Learning Analytics and Knowledge, ACM, 2011, pp.81-85.

[7] L. Lockyer and S. Dawson, "Learning designs and learning analytics", Proceedings of the 1st International Conference on Learning Analytics and Knowledge, ACM, 2011, pp.153-156.

[8] G.Richards and I. DeVries, "Revisiting formative evaluation: dynamic monitoring for the improvement of learning activity design and delivery", Proceedings of the 1st International Conference on Learning Analytics and Knowledge, ACM, 2011, pp.157-162.

[9] A. Pardo and C. D. Kloos, "Stepping out of the box: towards analytics outside the learning management system", Proceedings of the 1st International Conference on Learning Analytics and Knowledge, ACM, 2011, pp.163-167.

[10] P. Blikstein, "Using learning analytics to assess students' behavior in open-ended programming tasks", Proceedings of the 1st International Conference on Learning Analytics and Knowledge, ACM, 2011, pp.110-116.

[11] S. E. Fancsali, "Variable construction for predictive and causal modeling of online education data", Proceedings of the 1st International Conference on Learning Analytics and Knowledge, ACM, 2011.

[12] K. Verbert, H. Drachsler, N. Manouselis, M. Wolpers, R. Vuorikari, and E. Duval, "Dataset-driven research for improving recommender systems for learning", Proceedings of the 1st International Conference on Learning Analytics and Knowledge, ACM, 2011, pp. 44-53

[13] H. Drachsler, T. Bogers, R. Vuorikari, K. Verbert, E. Duval, N. Manouselis, G. Beham, S. Lindstaedt, H. Stern, M. Friedrich and M. Wolpers "Issues and considerations regarding sharable data sets for recommender systems in technology enhanced learning", Procedia Computer Science, 2010, 1, pp. 2849 - 2858.

[14] K. Verbert,N. Manouselis, H.Drachsler and E. Duval, "Dataset-Driven Research to Support Learning and Knowledge Analytics", Educational Technology & Society, 2012, 15, pp. 133-148.

[15] B. Taraghi, M. Ebner, G. Till and H. Mühlburger, "Personal Learning Environment-a Conceptual Study", iJET, International journal of emerging technologies in learning, 2010, 5, pp. 25-30.

[16] M.Ebner and B. Taraghi, "Personal Learning Environment for Higher Education--A First Prototype", World Conference on Educational Multimedia, Hypermedia and Telecommunications, 2010, pp. 1158-1166.

[17] B. Taraghi, M. Ebner and S. Schaffert "Personal learning environments for higher education: A mashup based widget concept", Proceedings of the Second International Workshop on Mashup Personal Learning Environments (MUPPLE09), Nice, France, 2009, 1613-0073.

[18] M. Ebner, N. Scerbakov, B. Taraghi, W. Nagler and I. Kamrat, "Teaching and Learning in Higher Education – An Integral Approach", Proceedings of Society for Information Technology & Teacher Education International Conference 2010, Gibson, D. & Dodge, B. (Eds.) , AACE , 2010 , pp. 428-436.

[19] B. Taraghi, M. Ebner and K. Clemens, "Personal Learning Environment -- Generation 2.0", World Conference on Educational Multimedia, Hypermedia and Telecommunications, 2012, pp. 1828-1835.

[20] B. Taraghi, C. Stickel and M. Ebner, "Survival of the Fittest -- Utilization of Natural Selection Mechanisms for Improving PLE", Proceedings of the first Workshop on Exploring the Fitness and Evolvability of Personal Learning Environments, 2011, pp. 4-9.

[21] V. Mulwad, T. Finin, Z. Syed and A. Joshi, "Using linked data to interpret tables", Proceedings of the the First International Workshop on Consuming Linked Data, 2010.

[22] Z. Jeremić, J. Jovanović and D. Gašević, "Personal learning environments on the Social Semantic Web Semantic Web", IOS Press, 2012.

[23] S. Softic, B. Taraghi and W. Halb, "Weaving Social E-Learning Platforms Into the Web of Linked Data", I-SEMANTICS, 2009, pp. 559-567.

Selver Softic obtained his master's degree with distinction in the field of computer science. Currently he is writing his PhD on topic of application of Linked Data for learning management systems supervised by Prof. Ebner at Graz University of Technology. Mr. Softic visited in year 2013 the Multimedia Lab (Semantic Web Task Force) led by Prof. Van De Walle at Ghent University as visiting researcher . He is author of various papers on PLEs, linked Data and Semantic Web.