

biblio.ugent.be

The UGent Institutional Repository is the electronic archiving and dissemination platform for all UGent research publications. Ghent University has implemented a mandate stipulating that all academic publications of UGent researchers should be deposited and archived in this repository. Except for items where current copyright restrictions apply, these papers are available in Open Access.

This item is the archived peer-reviewed author-version of:

Mining and Visualizing Usage of Educational Systems Using Linked Data

Selver Softic, Behnam Taraghi, Martin Ebner, Laurens De Vocht, Erik Mannens, and Rik Van de Walle

In: *Immersive Education*, 17–26, 2015.

http://dx.doi.org/10.1007/978-3-319-22017-8_2

To refer to or to cite this work, please use the citation to the published version:

Softic, S., Taraghi, B., Ebner, M., De Vocht, L., Mannens, E., and Van de Walle, R. (2015). Mining and Visualizing Usage of Educational Systems Using Linked Data. *Immersive Education* 17–26. [10.1007/978-3-319-22017-8_2](http://dx.doi.org/10.1007/978-3-319-22017-8_2)

Mining and Visualizing Usage of Educational Systems Using Linked Data

Selver Softic¹(✉), Behnam Taraghi¹, Martin Ebner¹, Laurens De Vocht², Erik Mannens², and Rik Van de Walle²

¹ Department for Social Learning, Graz University of Technology- IICM,
Muenzgraben street 35A, 8010 Graz, Austria
{selver.softic,b.taraghi,martin.ebner}@tugraz.at,
<http://www.tugraz.at>

² iMinds - Multimedia Lab, Ghent University,
Gaston Crommelaan 8, 9000 Ghent, Belgium
{laurens.devocht,erik.mannens,rik.vandewalle}@ugent.be
<http://www.ugent.be>

Abstract. This work introduces 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 improvements and monitoring such systems, in generally, with respect to social, functional, user and activity centric level [7,15]. The case study demonstrates the application of semantic modelling of the activity context, from data collected over two years from the PLE at Graz University of Technology, using adequate domain ontologies, semantic technologies and visualization as reflection for potential technical and functional improvements. As it will be shown, this approach offers easy interfacing and extensibility on technological level and fast insight on statistical and preference trends for analytic tasks.

Keywords: Data mining · Linked data · Micro content · Education · Research

1 Introduction

Modern learning environments, beside learning resources provided by the educational institution, aim at integration of popular internet services that might be of interest of learners like: Google Hangout, Facebook, YouTube, Newsgroups, Twitter, Slideshare just to name some of them. Maintaining such platforms is intensively changing process demanding from maintainers to actively adapt their systems to the learner needs. Nowadays, learners are expecting focused and simple platforms helping them to organise their learning process. Learners don't want to waste their time on informations and actions which could disturb or prolong their learning. Therefore user adaptivity is 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 offers new

opportunities for optimization [13]. 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 [9] useful for optimization of PLE and adapting them to the learners by using more personalization e.g. through recommendation of interesting learning widgets.

2 Related Work

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

2.1 Learning Analytics and Importance of Tracking and Reflection of User Logs

The current learning analytics research community defines [16] learning analytics as the analysis of communication logs [1, 15], learning resources [11], learning management system logs as well existing learning designs [8, 14] and the activity outside of the learning management systems [2, 12]. The result of this analysis improves the creation of predictive models [6, 18], recommendations [3, 25] and reflection [26]. 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. user behaviour tracking plays a decisive role for the evaluation. Different works outlined the importance of tracking activity data in Learning Management Systems [9, 15, 16, 25, 26]. 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.

2.2 PLE at Graz University of Technology

The main idea of PLE at Graz University of Technology¹ 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 [5, 23]. The TU Graz PLE contains widgets [5, 22, 23] 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 [4, 21]. 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 [24] (Fig. 1).

¹ <http://ple.tugraz.at>.

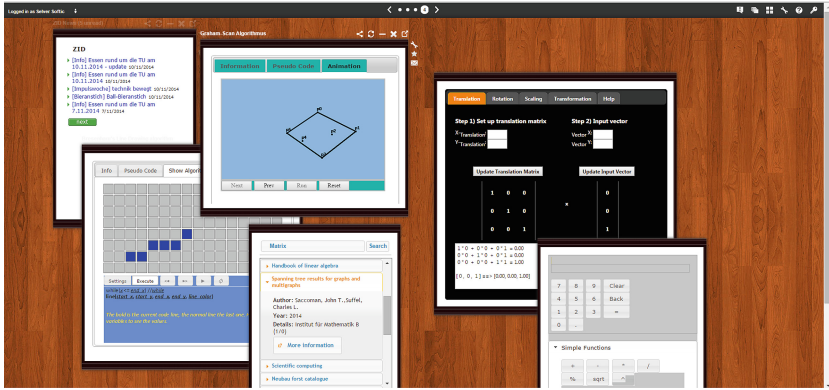


Fig. 1. PLE at Graz University of Technology

2.3 Semantic Modeling of Activities in PLE

The Semantic Web standards like RDF² and SPARQL³ enable data to be and for interchange 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 [10] to advance the learning process [7, 13] as well to support the controlled knowledge generation in E-learning environments [20]. This potential was also recognised by resent research in *IntelLEO Project*⁴. *IntelLEO* delivered an ontology framework where *Activities Ontology*⁵ is used to model learning activities and events related to them. Due to the relatedness to the problem that is addressed by this work this ontologies have been used to model the context of analytic data collected from PLE logs.

3 Approach for Mining Usage Logs

Presented approach is based on transforming collected data from PLE logs into instances of *Activities Ontology*. This process produces as output Linked Data graphs query able by SPARQL standard query language. The SPARQL is applied to query the Linked Data and mine the output for analytic visualizations (see Fig. 2). The overall goal of this processes is summarization, visualizations and evaluation of statistic data that enable the PLE optimization, in interface design and adaptation of content of PLE to the learner. This approach is inspired by the examples from current research in the area of Self-regulated Learners (SRL) [7, 19].

² <http://www.w3.org/RDF/>.

³ <http://www.w3.org/TR/rdf-sparql-query/>.

⁴ <http://intelleo.eu>.

⁵ <http://www.intelleo.eu/ontologies/activities/spec/>.

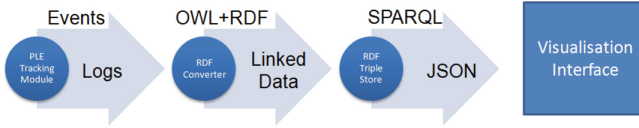


Fig. 2. Mining pipe line for PLE usage logs

3.1 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 then 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 [12, 17].

3.2 Modeling Usage Logs

The main precondition for meaningful mining of usage trends is choice of appropriate data model since RDF offers only the framework how structure and link data. This task concerns mostly the choice of the right vocabulary or ontology. *Activities Ontology* offers a vocabulary to represent different activities and events related to them inside of a learning environment with possibility to describe and reference the environment (in this case PLE) where these activities occur. Formulation in Listing 1.1 depicts an instance of usage AO:LOGGING instance. This excerpt comes from the tracking module. Such data is stored in a memory RDF Store (Graph Database for Linked Data) with SPARQL Endpoint (interface where Linked Data can be queried). This sample instance reflects that a usage AO:LOGGING event occurred at certain time point inside the learning widget named *LatexFormulaToPngWidget* as AO:ENVIROMENT. As shown in this example vocabularies and ontologies which suits well to specific case enriches the analytic process with a high level of expressiveness in a very compact manner.

3.3 Querying Usage Logs

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 15 used widgets?”. Listing 1.2 represents exactly this question stated in the manner of SPARQL syntax.

4 Preliminary Results, Conclusion and Outlook

Advantages of Linked Data approach is usage of standardized web technologies which are scalable and flexible regarding the changes of representation structure

Listing 1.1. Sample model of a log for a PLE widget in N3 notation

```

@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#> .

<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:Environment;
  rdfs:label "LaTeXFormulaPNG Converter" .

```

of data. Also very important aspect of mining PLE usage data using Linked Data is for sure high operational tolerance regarding incomplete analytic data instances as well as easier interfacing to other systems which could make use of information provided by PLE.

SPARQL as query language which operates over the Linked Data graphs of usage logs offers much flexibility regarding the generation of results, in different state of the art output formats, that should be visualized in end instance. It also allows on-demand statistical cumulations that can be used in the future as basic stats for recommendation of new widgets in the PLE or similar tasks.

Listing 1.2. Querying the intensity of usage of top 15 widgets in PLE.

```

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#> .

SELECT DISTINCT ?widgetname ?date (COUNT(?widgetname) AS ?count)
WHERE
{
  ?x rdf:type ao:Logging;
    ao:occursIn ?widget;
    ao:timestamp ?date.

  ?widget rdf:type ao:Environment;
    rdfs:label ?widgetname.
}
GROUP BY ?widgetname
ORDER BY DESC(?count)
LIMIT 15

```

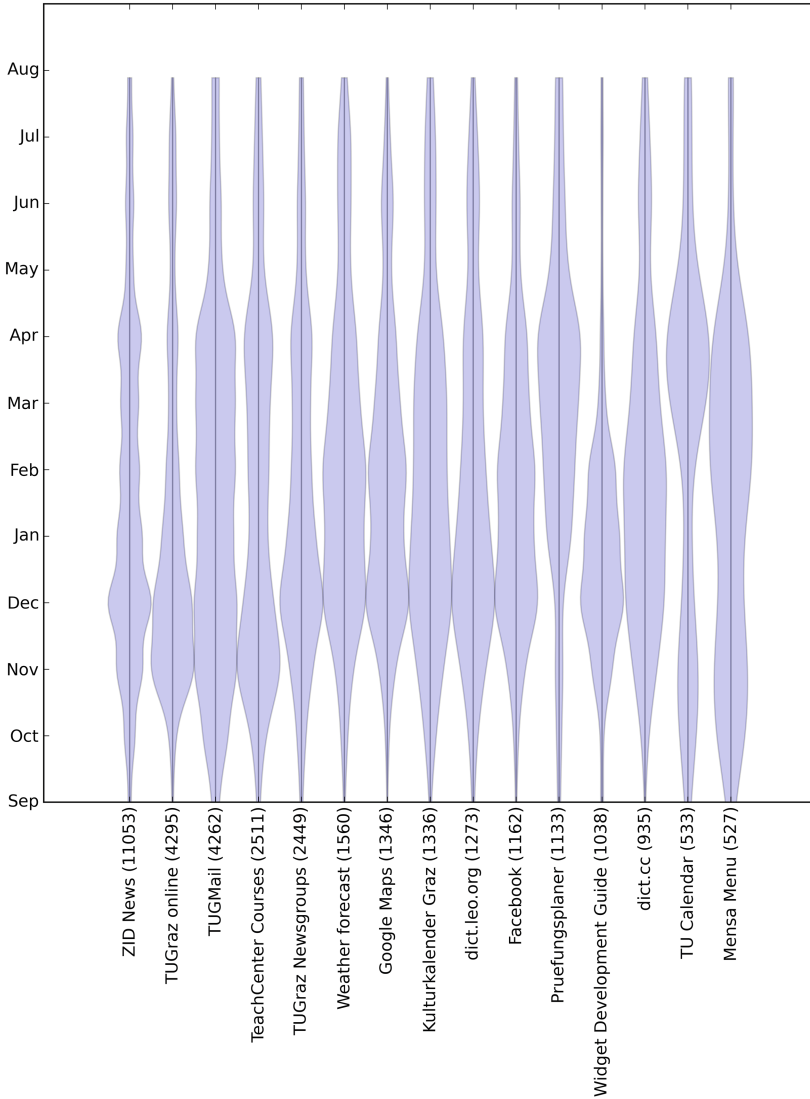


Fig. 3. Top 15 activities for an academic year time period 2011-2012

As preliminary result presented approach allows us mining the trends of PLE widgets usage overall time periods like presented in Fig. 3. This violin graph depicts the visual answer of the query from Listing 1.2. Also the intensity shows that as expected that most activity on widgets happens at the beginning when PLE is presented in introductory lectures to the newcomers and freshmen and at the end of academic terms when most of the students prepare for examinations. The statistics visualisation help us to gain deep insight into the behaviour of a users in a certain period of time. Presented approach generates uniform interfaces

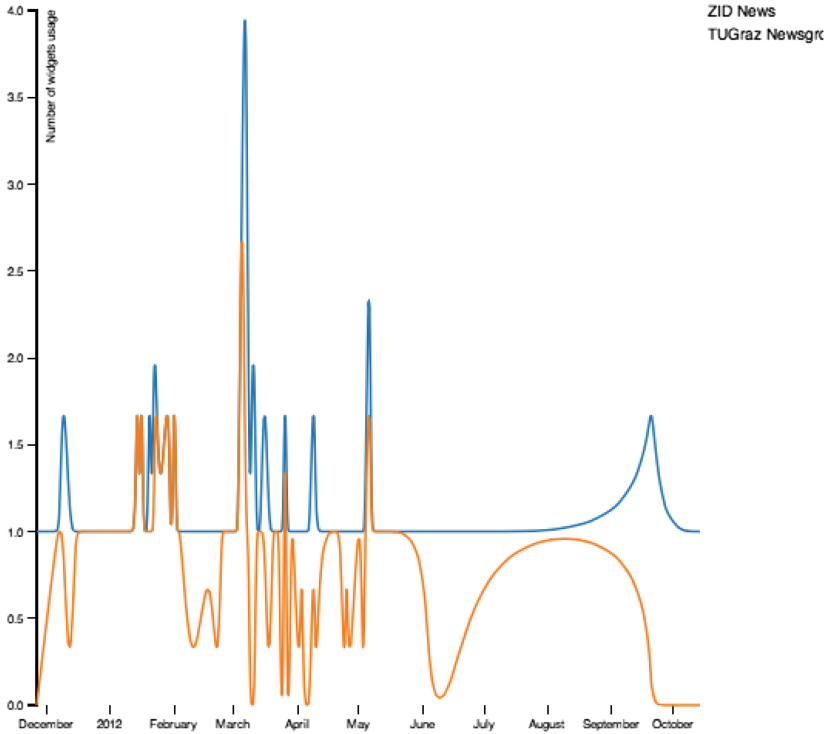


Fig. 4. Comparison of two news widgets.

for information exchange, enables flexibility for visual analytics, and also includes the flexibility regarding the enrichment of learning analytics data with Linked Data sources from the Web. The spread of applicability covers wide range of analytics methodologies like prediction, reflection and as result of these the intervention field. Figure 4 reflects the advantage of such approach where e.g. two widgets with similar purpose can be visually compared (in this case two news-groups widgets). Future efforts regarding improvement semantic structure data layer, besides the mentioned Linked Data could also include precisely defined categorisation, userwise 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 Fig. 5. By tracking the usages on user level the teachers will be able to draw conclusions about the popularity and quality of their learning widgets.

The overview over distribution of usage logs can reflect the overall interest of the users within PLE. Such inputs evaluated and interpreted in appropriate way contribute implicitly the improvement of the quality of services for students and teachers. The PLE becomes, in technical manner, extensible and well connected by standardized and intelligent interfaces and available for other web based tools and services.

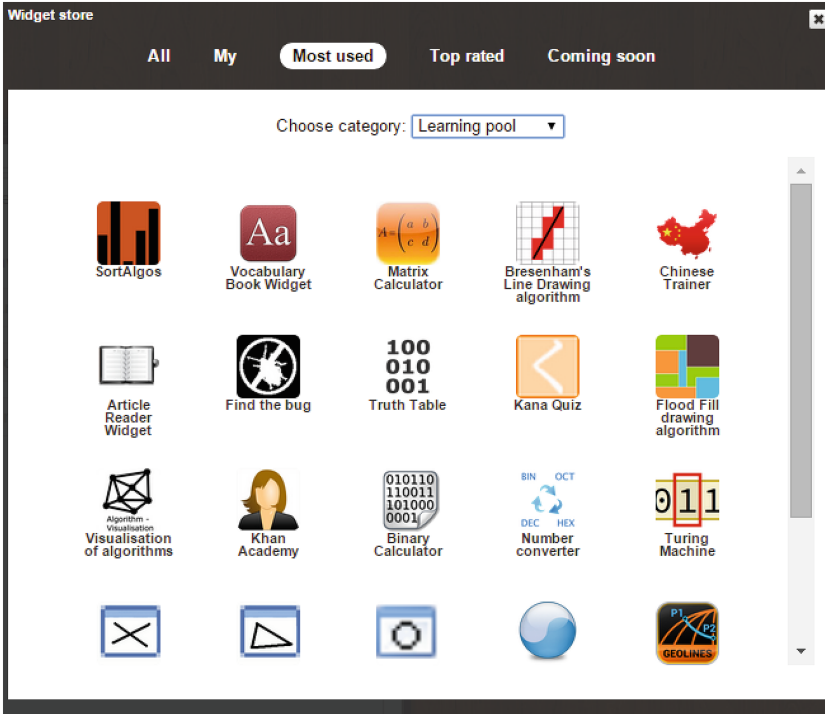


Fig. 5. PLE Widget store recommendations based upon usage log statistics.

Acknowledgements. 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.

References

1. Bakharia, A., Dawson, S.: Snapp: a bird's-eye view of temporal participant interaction. In: Proceedings of the 1st International Conference on Learning Analytics and Knowledge, pp. 168–173. LAK 2011, ACM, New York, NY, USA (2011). <http://doi.acm.org/10.1145/2090116.2090144>
2. Blikstein, P.: Using learning analytics to assess students' behavior in open-ended programming tasks. In: Proceedings of the 1st International Conference on Learning Analytics and Knowledge, pp. 110–116. LAK 2011, ACM, New York, NY, USA (2011). <http://doi.acm.org/10.1145/2090116.2090132>

3. Drachsler, H., Bogers, T., Vuorikari, R., Verbert, K., Duval, E., Manouselis, N., Beham, G., Lindstaedt, S., Stern, H., Friedrich, M., Wolpers, M.: Issues and considerations regarding sharable data sets for recommender systems in technology enhanced learning. *Procedia Comput. Sci.* **1**(2), 2849–2858 (2010). <http://www.sciencedirect.com/science/article/pii/S1877050910003236>, proceedings of the 1st Workshop on Recommender Systems for Technology Enhanced Learning (RecSysTEL 2010)
4. Ebner, M., Scerbakov, N., Taraghi, B., Nagler, W., Kamrat, I.: Teaching and learning in higher education an integral approach. In: Gibson, D., Dodge, B. (eds.) *Proceedings of Society for Information Technology and Teacher Education International Conference 2010*, pp. 428–436. AACE, San Diego (2010). <http://www.editlib.org/p/33375>
5. Ebner, M., Taraghi, B.: Personal learning environment for higher education—a first prototype. In: *World Conference on Educational Multimedia, Hypermedia and Telecommunications*, pp. 1158–1166 (2010)
6. Fancsali, S.E.: Variable construction for predictive and causal modeling of online education data. In: *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, pp. 54–63. LAK 2011, ACM, New York, NY, USA (2011). <http://doi.acm.org/10.1145/2090116.2090123>
7. Jeremić, Z., Jovanović, J., Gašević, D.: Personal learning environments on the social semantic web. *Semant. Web* **2**(12), 8–12 (2012)
8. Lockyer, L., Dawson, S.: Learning designs and learning analytics. In: *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, pp. 153–156. LAK 2011, ACM, New York, NY, USA (2011). <http://doi.acm.org/10.1145/2090116.2090140>
9. Mazza, R., Milani, C.: Exploring usage analysis in learning systems: gaining insights from visualisations. In: *Workshop on Usage Analysis in Learning Systems, Proceedings of Artificial Intelligence in Education*. Amsterdam (2005)
10. Mulwad, V., Finin, T., Syed, Z., Joshi, A.: Using linked data to interpret tables. In: *Proceedings of the the First International Workshop on Consuming Linked Data*, November 2010
11. Niemann, K., Schmitz, H.C., Scheffel, M., Wolpers, M.: Usage contexts for object similarity: exploratory investigations. In: *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, pp. 81–85. LAK 2011, ACM, New York, NY, USA (2011). <http://doi.acm.org/10.1145/2090116.2090127>
12. Pardo, A., Kloos, C.D.: Stepping out of the box: towards analytics outside the learning management system. In: *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, pp. 163–167. LAK 2011, ACM, New York, NY, USA (2011). <http://doi.acm.org/10.1145/2090116.2090142>
13. Prinsloo, P., Slade, S., Galpin, F.: Learning analytics: challenges, paradoxes and opportunities for mega open distance learning institutions. In: *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, pp. 130–133. LAK 2012, ACM, New York, NY, USA (2012). <http://doi.acm.org/10.1145/2330601.2330635>
14. Richards, G., DeVries, I.: Revisiting formative evaluation: dynamic monitoring for the improvement of learning activity design and delivery. In: *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, pp. 157–162. LAK 2011, ACM, New York, NY, USA (2011). <http://doi.acm.org/10.1145/2090116.2090141>

15. Rosen, D., Miagkikh, V., Suthers, D.: Social and semantic network analysis of chat logs. In: Proceedings of the 1st International Conference on Learning Analytics and Knowledge, pp. 134–139. LAK 2011, ACM, New York, NY, USA (2011). <http://doi.acm.org/10.1145/2090116.2090137>
16. Santos, J.L., Govaerts, S., Verbert, K., Duval, E.: Goal-oriented visualizations of activity tracking: a case study with engineering students. In: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, pp. 143–152. LAK 2012, ACM, New York, NY, USA (2012). <http://doi.acm.org/10.1145/2330601.2330639>
17. Santos Odriozola, J.L., Verbert, K., Govaerts, S., Duval, E.: Visualizing PLE usage. In: Proceedings of EFEPLE11 1st Workshop on Exploring the Fitness and Evolvability of Personal Learning Environments, pp. 34–38. CEUR WS, Aug 2011. <https://lirias.kuleuven.be/handle/123456789/322006>
18. Sharkey, M.: Academic analytics landscape at the university of phoenix. In: Proceedings of the 1st International Conference on Learning Analytics and Knowledge, pp. 122–126. LAK 2011, ACM, New York, NY, USA (2011). <http://doi.acm.org/10.1145/2090116.2090135>
19. Siadaty, M., Jovanovic, J., Pata, K., Holocher-Ertl, T., Gasevic, D., Milikic, N.: A semantic web-enabled tool for self-regulated learning in the workplace. In: ICALT, pp. 66–70. IEEE Computer Society (2011). <http://dblp.uni-trier.de/db/conf/icalt/icalt2011.html#SiadatyJPHGM11>
20. Softic, S., Taraghi, B., Halb, W.: Weaving social e-learning platforms into the web of linked data. In: I-Semantics, pp. 559–567 (2009)
21. Taraghi, B., Ebner, M., Clemens, K.: Personal learning environment - generation 2.0. In: World Conference on Educational Multimedia, Hypermedia and Telecommunications, pp. 1828–1835. AACE (2012)
22. Taraghi, B., Ebner, M., Schaffert, S.: Personal learning environments for higher education: a mashup based widget concept. In: Proceedings of the Second International Workshop on Mashup Personal Learning Environments (MUPPLE 2009), Nice, France, pp. 1613–0073 (2009)
23. Taraghi, B., Ebner, M., Till, G., Mühlburger, H.: Personal learning environment-a conceptual study. *iJET Int. J. Emerg. Technol. Learn.* **5**(S1), 25–30 (2010)
24. Taraghi, B., Stickel, C., Ebner, M.: Survival of the fittest - utilization of natural selection mechanisms for improving ple. In: Proceedings of the First Workshop on Exploring the Fitness and Evolvability of Personal Learning Environments, pp. 4–9 (2011)
25. Verbert, K., Drachsler, H., Manouselis, N., Wolpers, M., Vuorikari, R., Duval, E.: Dataset-driven research for improving recommender systems for learning. In: Proceedings of the 1st International Conference on Learning Analytics and Knowledge, pp. 44–53. LAK 2011, ACM, New York, NY, USA (2011). <http://doi.acm.org/10.1145/2090116.2090122>
26. Verbert, K., Manouselis, N., Drachsler, H., Duval, E.: Dataset-driven research to support learning and knowledge analytics. *Educ. Technol. Soc.* **15**(3), 133–148 (2012)