

Visual learning analytics techniques applied in software engineering subjects

Miguel Á. Conde

Department. of Mechanics, Computer Science and
Aerospace Engineering.
University of Leon
León, Spain
miguel.conde@unileon.es

Francisco J. García-Peñalvo, Diego A. Gómez-
Aguilar, Roberto Theron

GRIAL Research Group, Department of Computers and
Automatics, Research Institute for Educational Sciences,
University of Salamanca
Salamanca, Spain
{fgarcia,dialgoag,theron}@usal.es

Abstract— The technology applied to educational contexts, and specially the learning platforms, provides students and teachers with a set of tools and spaces to carry out the learning processes. Information related to the participation and interaction of these stakeholders with their peers and with the platform is recorded. It would be useful to exploit this information in order to make decisions. However this is a complex activity mainly because of the huge quantity of information stored. This work presents a visual learning analytics system that makes possible the exploitation of that information. The system includes several tools that help to analyze users' interaction attending to different dimensions, such as: when interaction is carried out, which are more important contents for users, how they interact with others, etc. This system has been tested with the subject information recorded during five academic years. From this analysis it is possible to show that visual learning analytics may help to improve educational practices.

Keywords—*Visual Analytics, Learning Analytics, LMS, Data Mining*

I. INTRODUCTION

Information and Communication Technologies (ICT) have been applied to different educational areas. Clear examples of this are Learning Management Systems (LMSs). They are very popular and common solutions both in industry and academic contexts [1]. LMSs are especially useful for learners and teachers. These environments are focused on course delivery and provide teachers with tools that not only support but also extend the traditional concept of the classroom and facilitate managerial tasks [2]. These systems also provide students with spaces in which they may perform their academic activities, complement their lectures and, to a greater or lesser extent, collaborate with other students and teachers. This means that LMSs are platforms that support learning/teaching processes. However they do not only provide functionalities to be employed in such processes but also gather information about the interaction of their users and the events they carry out on them. If this data could be exploited it would be possible to know what has happened and is happening in the LMS. However the main problem of this information is that it is stored as raw data. This means a huge quantity of complex

information that should be analysed to obtain real meaning from it.

The exploitation of the gathered data is not an easy task. For instance, the Software Engineering subject of the Computer Science Degree of the University of Salamanca has around 50000 records per year. With these records it is possible to explore if the subject is progressing as expected, how are interacting students with peers and teachers, if teachers are participating or not in forums, which are the concepts more discussed in forums, if the amount of interaction is linked to better grades, etc. But to do so it is necessary to present all the stored information in a comprehensive way, so the people in charge of a learning activity can identify the most relevant aspects and make decisions from them. Learning Analytics (LA), Academic Analytics (AA) or Educational Data Mining (EDM) can be used to do this.

According to Buckingham Shum and Ferguson, the earliest mention of the term Learning Analytics dates back to 2000 [3]. But that mention was marginal and, in the beginning, Learning Analytics was a derivation of the principles of Business Intelligence and Data Mining used in Enterprise Information Systems applied to the educational field, in a very theoretical and prospective way [4]. However with the advent and worldwide expansion of LMSs and Massive Online Courses (MOOCs) more quantity of information is stored [5].

Given this context, there are several definitions about Learning Analytics and not all authors agree on them. A possible one was the posed by Ferguson “the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs [6]”.

Other possible trend to apply is Academic Analytics. It was defined by Goldstein and Katz [7] as the application of business intelligence tools to learning. Academic Analytics aims to go beyond the information report to facilitate decision-making. These decisions are mainly focused on the institution and the organization while Learning Analytics is more focused on making decisions related to students, teachers and the learning process itself [8].

Educational Data Mining is another possibility. It is a field that exploits statistical, machine-learning, and data-mining algorithms over the different types of educational data. Its main objective is to analyse these types of data in order to resolve educational research issues [9]. Learning Analytics also shares many attributes with Educational Data Mining [10]. Thus, whether Educational Data Mining is a subset of Learning Analytics or a different discipline with some points in common is for many still an open debate [4].

Despite the differences between Learning Analytics, Educational Analytics and Academic Analytics, they have a common objective that is the improvement and better understanding of teaching and learning processes. In order to achieve this, it is necessary to facilitate common ways to visualize and exploit the results of the analysis. In this case the idea is to explore logs stored in LMSs in order to make decisions to improve learning and teaching of a subject. This improvement can take place while the subjects are being developed or once they have finished. In order to do so analyzed information should be represented in a way that facilitates making decisions [11] and in this sense Visual Analytics has an important role.

This term was coined by Thomas and Cook [12] as “the science of analytical reasoning facilitated by interactive visual interfaces”. It “combines automated analysis techniques with interactive visualizations to create an effective understanding, reasoning and decision making process, based on very large and complex data sets” [13].

The present work describes how are exploited the Software Engineering records from University of Salamanca LMS but in a way that not only facilitates visualizing the information, but also interacting with it. That is, teachers can filter, search, or change dynamically the representation of the information. To do so a Visual eLearning Analytics (VeLA) system is used. It includes four main functionalities: 1) A Semantic Spiral Timeline that facilitates tracking the activity during specific periods of time; 2) A Semantic Tagcloud that allows the teacher to find the most important words in forums filtered by different parameters; 3) A social graph that allows distinguishing the interactions among the students and teachers in different subjects; and 4) A tool to compare and establish relationships among the data stored in the LMS and the activities the users carry out on them. These tools are tested with the logs of Software Engineering Subject during five years, and it is possible to extract knowledge from them that facilitates making decisions.

The paper is structured as follows. Section 2 compares VeLA with other related works. Section 3 describes the theoretical model that supports the visual analytics tools. Section 4 shows the tools integrated by the visual analytics system. Section 5 presents a case study with the information of a Software Engineering subject. Finally, in Section 6, some conclusions are posed.

II. RELATED WORKS

Learning analytics is a technique that has become quite important since ICT application to learning environments [14]. There are a wide set of Learning Analytics tools that can help

making decisions about teaching and learning processes. Taking into account the tools applied to LMS and attending to their purpose they can be classified in:

- Dashboards with general propose that can be applied to different platforms [15, 16], or to a specific one [17]. They use to provide information about students or teachers activity in the platform as information presented in tables and enrich with some bar charts or pie charts. The main problem of these solutions is that they are focused on the description of the activity carried out in LMS using very specific metrics and do not providing relationships between the information.
- Ad-hoc defined tools. These tools are defined to track or analyse very specific information and for a specific context. The problem of these solutions is that they are neither flexible nor scalable. Examples of these tools are [18-22].
- Learning Analytics tools to analyse specific issues. These tools can be applied to different platforms and use to be focused on specific information with a specific type of representation. The main drawback of these approaches is that the analysis they carry out can fit or not with the necessities of an institution. Some examples can be LEMO [23], SNAPP [24], StepUp![25], etc.
- Learning Analytic framework or tools. They can be applied to several platforms or contexts and explore different issues through different visual representations. Tools such as SAM [26] or VeLa [27] can be included in this group.

Regarding with the tools applied to Moodle, VeLA is not only one. Other examples can be: GISMO Moodle, a dashboard that gathers log information and shows some graphical representation [28]; Moodle Dashboard [29], other tool with several ways to represent log information; Moodog which visualizes metrics of activity logs in Moodle using bar charts [30]; SNAPP, a social interaction learning analytics tool that can be applied to several LMS [24]; Google Analytics plugin [16]; etc.

VeLA system goes beyond other LA tools because:

- VeLA is a LA framework supported by visual analytics techniques. Comparing with other initiatives, especially dashboards, that only present information in tables and some bar charts or pie charts [31], VeLA allows interaction with the dataset and presents linked views of the data. For instance, the user can select a specific period of time in the spiral view, immediately the changes made will affect to both the tagcloud and the social graph views (see section 4).
- VeLA can be applied to several learning platforms because it uses web services to interact with the LMS, so it is possible to easily be adapted to other platforms.

In the literature the tools more similar to VeLA are those included in the last category of the previously described.

III. VISUALIZATION MODEL FOR EDUCATIONAL INFORMATION

This work tries to define tools and processes that make easier the analysis of learning evidences. In this way it would be possible to make decisions that help to improve students' learning and that can be applied to subjects such as Software Engineering.

The present section describes some of the research that supported the visual analytics model, the process to articulate this model and the main issues that it should take into account to be applied to eLearning courses.

A. The teoretical model

The aim of a visual analytics system is to facilitate the exploitation of information stored in educational contexts to improve learning processes. To do so visual analytic should join the advantages of different analytics trends such as LA, AA or EDM.

Visual Analytics application to education provides highly interactive tools that show data from different perspectives. These tools facilitate both the exploration of the information and the confirmation of hypothesis about educational issues.

In this case the visualization model to use takes into account previous research related to the above mentioned analytics trends such as: Chatti et al. [32] LA reference model; Clow LA cicle theory [33]; Keim [13] visual data exploration and some issues and constraints from Greller and Draschsler critical dimensions model [34].

Therefore the main contribution of this model is the combination of automatic analytic and exploratory techniques to facilitate information understanding as well as the interaction with that information.

The model is articulated through a process based on the work of Ben Shneiderman [35] and extended by Daniel Keim, et al. [36]. The idea is to facilitate interactive visualizations with which the users can explore the information to extract abstract models from sets of data too big or too complex to be easily analysed. The process can be understood as an iterative sequence of stages that take into account different stakeholders such as students, teachers, people in charge of institutions and other staff. These stages are:

- Requirements analysis. During this stage it is necessary establish what information to consider, how it is recorded, which are the constraints of this analysis (if any), which are the expected results, etc.
- Data Model. During this stage is defined what data should be considered and how it is described.
- Representation Model. This stage involves two states, visualization and the definition of analytic models. Visualization provides the users a way to represent and interact with the information while the definition of analytic models is related to what techniques are used to analyse the data and which parameters can be manipulated in the representations. Both states are related, analytics models can be modified depending on the kind of representation; and the way in which

information is analysed can change how data and interaction are presented.

- Knowledge Management and Process Feedback. After the representation of the information it is possible to obtain knowledge about learning processes and about the analytic system. This information can be used to improve the visual analysis process and in future iterations and/or in similar situations.

B. Visual analytics system for eLearning

With the model described in the previous section it is necessary to consider which are most important factors that can be analyzed in educational contexts.

As for this work LMSs and online courses are studied, the visual analytic system should take into account the following issues: 1) eLearning processes may not be restricted to specific periods of time; 2) courses and subjects length can depend on the context where they are carried out (it is not the same a degree subject than an specialization course related to a institution); 3) eLearning courses have some aspects to take specially into account such as the interaction of the teachers and students during the courses, the quality of their participation, etc.

Given this context the Visual Analytics system should consider how interaction evolve over the time; what contents are relevant for students and what are not; which are most active users in forums; which is the relationship between participation and students' performance; etc.

These issues are addressed by the Visual Analytics System through the development of four tools that facilitates the analysis of courses, subjects and campus during the time; the analysis of content by using a sematic tag-cloud; the analysis of interaction through social network representation; and the analysis of frequency of different category of interactions.

Next section describes how the visual analytics system is implemented and the main functionalities of the tools that it includes.

IV. VE LA SYSTEM

In order to apply the visual analytics model it is necessary to develop a set of tools. With this aim VeLA system was developed.

VeLA system is a Java visualization framework that includes 4 tools to explore different perspectives of the data stored in the LMS. These perspectives can be used separately or combined. For instance if a period of time is selected in one the tools the rest show also data according to that selection.

The current development of VeLA works for Moodle because of its popularity and because this LMS includes a web service layer that open it both to integration of tools and for the exploitation of information [37].

This last issue is very important in this case, because it facilitates accessing to learners' and teachers' activity stored in the LMS [38]. The use of web services to access to the LMS makes VeLA system independent of the programming language in which the LMS is implemented and independent of

a specific version of the selected LMS. VeLA can be used with Moodle but also with other LMSs with just a change in the set of web services used to access to the information. Most of the LMSs include a web service layer [39] so this guarantees VeLA portability.

With the information gathered from the LMS it is possible to show different visualizations that support decision-making processes.

A. Spiral Timeline

One of the issues that VeLA model should satisfy is to represent how the activity of learners and teachers evolves along the time. The system should consider different contexts, users, periods of time, etc.

In order to do so a Semantic Spiral Timeline (SST) is used [40]. The SST provides compact representation of the overall use of the LMS. The information view can be adapted to the user's requirements, so he/she can explore all the available temporal data, going from the overview to the detail of a given person or activity within a period of time. SST includes a visual technique to balance the detail and context in data visualization, known as semantic zooming or multi-scale interfaces. It changes the type and meaning of information displayed by the object [41].

The visualization has three panels (Fig 2) with different data views. The main representation is the spiral timeline, which, in its simplest form, is merely a sequence of color-coded events. These are ordered clockwise with the oldest data at the center of the spiral and the outermost data depicting the most recent event. The remaining two panels provide additional views of the data shown in the spiral. In the top view the temporal data is drawn in a linear representation, showing the current data on focus. Finally, on the right hand side, you find the data overview (i.e., the context), which also includes a slider for selecting the period in focus. The tool also includes the possibility to configure the parameters of the representation making selections, defining periods of time, etc. This is done through a contextual menu.

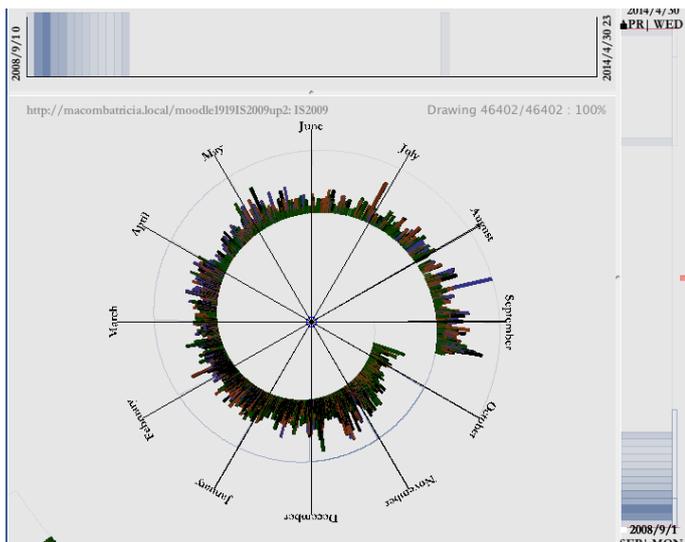


Fig. 1. SST representation.

The views presented by these three panels are inter-linked. This means that, since all the views are different ways of conveying the same information, when the user interacts with one of them and changes the representation, the other views are also changed depending on the action originally performed by the user and on how that action affects the shared dataset [40].

With these tools it is possible to find cyclic patterns in a selected period of time using the spiral.

B. Semantic Tagcloud

Other of the functionalities that VeLA provides is a sematic analysis of contents published by students and teachers. To do so a Semantic Tag-cloud (STC) was developed [42].

A tag-cloud usually has a particular purpose: to present a visual overview of a collection of text. It presents a certain number of most often used tags in a defined area of the user interface. A tag's popularity is expressed by its font size (relative to the other tags) and is therefore easily recognized [43]. Tag-clouds are also navigation interfaces as the tags are usually hyperlinks leading to a collection of items they are associated with.

In this case the tag-cloud facilitates the representation of students and teachers contributions in forums (Fig. 2). It uses different size for words depending on the frequency and facilitating the navigation through posts, forums, threads, for a person, a course or a defined selection.

In addition the tag-cloud is representing also the evolution of the tags over the time. Tag clouds can evolve as the associated data source changes. This desire to study trends and understand how text content or topics evolve over time has been the purpose of other visualizations such as the commonly used line graphs and bar charts [44]. The STC represents a wave graph and bar graph per each tag in order to show how it evolves over the time. This is shown in Fig 2., the bar-graph (as curves in the background of the word), the wave-graph (as lines that grow up from the bottom of the word and representation of the tag. To further improve readability of the curves and the bars, different colour can be assigned to then depending on the type of activity (if the tag is created, read or updated). In these graphs the x-axis encodes time and the y-axis encodes the significance of the word clouds.



Fig. 2. Semantic Tagcloud representing the words for a Software Engineering subject

In addition to these functionalities the tag-cloud tool included in VeLA provides a physical and a semantic zoom. A physical zoom, on the one hand, changes the size and visible detail of objects. A semantic zoom, on the other hand, changes the type and meaning of information displayed by the object [45]. When the users use the semantic zoom, depending on the context, the tool selects for analyse the forum-posts related with the word zoomed. For example, if the user makes double click over a course, forum or discussion, the tool uses for this analysis all forum posts from that course or discussion or forum. In the case of the user makes semantic zoom on a word, the tool takes all forum posts that have the specific word selected to do the reconstruction of the temporal tag-cloud.

Furthermore the user can choose what he wants to see. He can show or hide every element of the representation on the visualization.

With this tool it is possible to analyse if the forums are used properly or not, which are the more discussed terms, when they have been discussed, etc.

C. Social Networks Graph (SNG)

In LMSs interactions between students and teachers and their use of resources represent the participation in the platform. This participation can be assessed [46] and can be related to other indicators such as grades.

Though the users of these LMS have constructed massive graph structures of social connectivity, these connections are not properly represented. LMS merely shows the network connections of single individuals as a list. Articulated connections between students and teachers in these systems are not clear. Higher-level community patterns can be even harder to discern. This has problematic implications for members' ability to explore their online community and gauge both the scale and the individuals to which their self-reported personal information is exposed [47].

The SNG is used to represent the map of the relationships, with the links, and frequency of activities of the students and teachers, with the size of icons, in the same graph. On the graph, entities such as Person, Course, discussion, posts and resources (coded as shown in Fig. 3) are shown as nodes and how they interact are represented as links between them.

Graph element position is based on a physics simulation of interacting forces; by default, nodes repel each other, edges act as springs, and drag forces (similar to air resistance) are applied. This force between nodes depends of the weight of the elements, which depends of the number of relations that each of them has and its edge size depends on the element degree and its hierarchy depth. Node size is directly proportional to node degree (number of connections with other node). Also the social graph can be transformed depending of the distance - bacon number, "Degrees of Separation", DOS [48] - between the selected node and the others. The filter of separation number can be customized. The different nodes can be hidden or be colored individually or by groups in order to distinguish or consider only some type of interactions. For instance, teachers can be represented with a color, students with other, etc. Moreover the SNG includes textual searching tools.

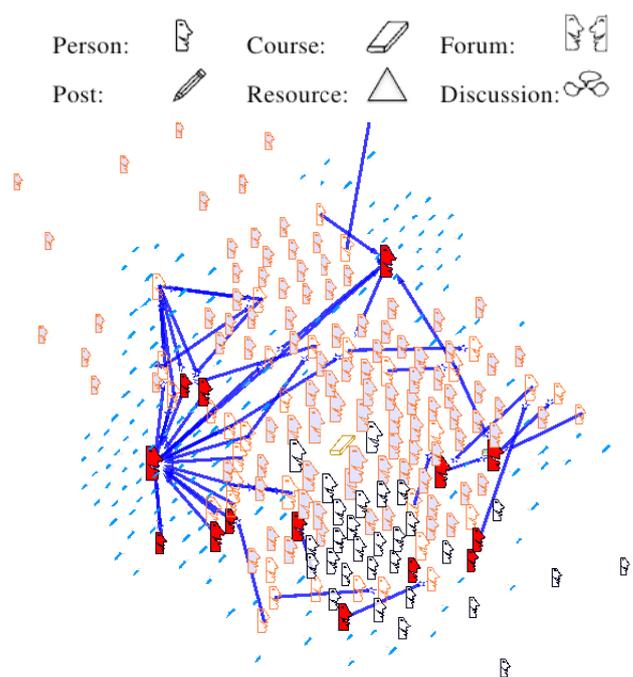


Fig. 3. SGN and legend

Also SNG makes possible to detect the frequency in the interaction of students in the forum and with the resources of a course, something that several studies relates to students' performance [46, 49].

D. Parallel coordinates representation

An important issue in the analysis of the interaction in learning platforms are the relationships between indicators. This can be done by using tools that facilitate the comparison between users activity attending to dimensions such as time, the forum, the role of the user, etc. In order to understand how different issues are related VeLA includes a specific tool. It is an interactive multidimensional visualization that uses a parallel coordinates technique [50].

This technique has been commonly used to show multivalued data. Dimensions are represented in vertical parallel axis and the data are represented as lines that connect these axis. In this case the dimensions (factors to compare) are represented as the horizontal axis, and the polygonal lines that connect them are the students' interaction. For instance some dimensions to consider how the participation is related with the grades could be the logs of the users, type of interaction, role of the user, action he/she carry out, module affected the specific user and grade (Fig 4).

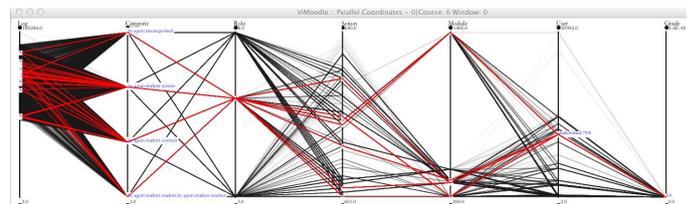


Fig. 4. Parallel coordinates representation to show how the activities carried out by a specific set of students affect the grades.

With this kind of representation it is possible to see which are the kind of activities that affect the grades, if the moment in which they are carried out is important, etc. In addition the tool facilitates the interaction with the representation by tagging spaces with more interaction (more lines), establishing groups of users to analyse (colouring their activities) and facilitating the definition of groups of activities.

V. CASE STUDY

In order to check how these visual analytics tools can be used and the knowledge they gather about students, teachers and their interaction, a case study is carried out. It is applied to a Software Engineering subject of the University of Salamanca. This section describes the subject and the outcomes obtained from the application of the tool.

A. Software Engineering Subject

Software Engineering is a discipline extremely important in the curriculum of the future Computer Science engineers. It comprises 60 hours (45 theoretical hours and 15 lab hours) and is focused on the following topics: Lifecycle and process requirement elicitation and documentation; Analysis and design methods and notation; and Modularity, software architecture, and software reuse principles.

During the practical part of the subject, students have to develop a small project, at the analysis and design levels, which they must do in groups of three members and which they have to present at the end of the semester. But also 10 hours of the time assigned to practice are reserved for organizing two-hour problem-solving workshops. Each workshop is devoted to a concrete modelling technique, and the students have the opportunity to discuss the solution to a given problem, so that they can learn from the mistakes and the corrections made during the workshop.

The Software Engineering subject is supported by the Moodle LMS of the University of Salamanca. The subject uses Moodle as a resource repository and a place to facilitate students' interaction. The information gathered by this LMS is exploited in order to discover necessities of the users, behaviour patterns of the students, new ways to improve the subject and so on.

The case study explores Software Engineering subject logs from 2008 to 2013, that is, five academic years. Each academic year has associated an average of 43000 logs and 160 students. This means around 21500 log entries and 800 students to be analysed for the five years period. In all these courses the subject has been developed in the first semester of the academic year (September to February).

B. Results of the application of the SST

SST is used to check when users access to the platform and which are the most critical moments for them. Theoretically, the increase should be focused on October, November and a part of December. The main reason is due to these are the busiest months and also when workshops take place. Data analysis results can be seen in Fig. 5. It shows that periods of increased activity correspond to the months of classroom teaching, with particularly representative peaks around workshops dates.

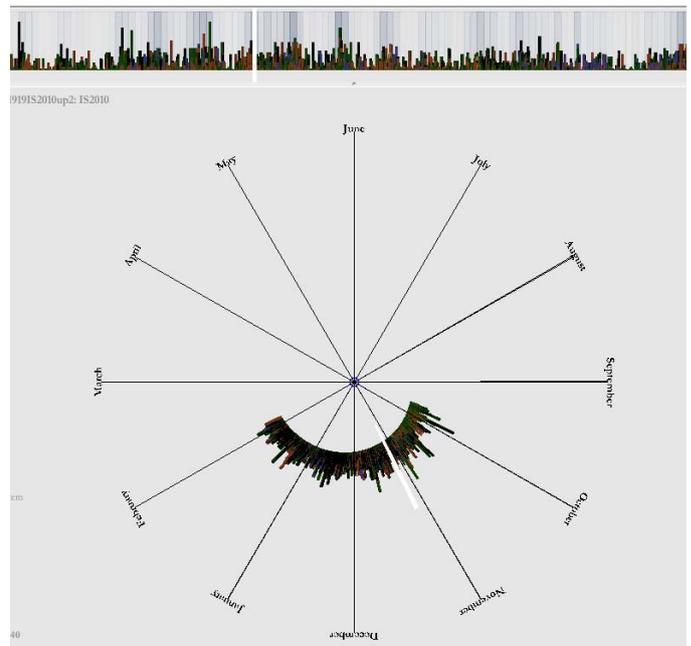


Fig. 5. Spiral Timeline to represent the activity of the first 4 months of the subject.

This makes sense since the students have to use the platform for their solution proposals delivery and because these exercises will be commented in the forums later. In addition it is possible to see an initial peak that represents the start of the course when the students begin to access to the contents published in the platform. During this period it can be interesting to introduce a greater presence of teachers. On the other hand it also would be interesting to increase students' participation during periods of low activity.

Moreover the SST gives teachers and idea about when during the week the accesses to resources and activities are more intensive (Fig. 6). In the case study the days are more relevant to students are Monday, Tuesday and Wednesday; teachers can use that information in order to introduce new resources in the subject when students usually read it.

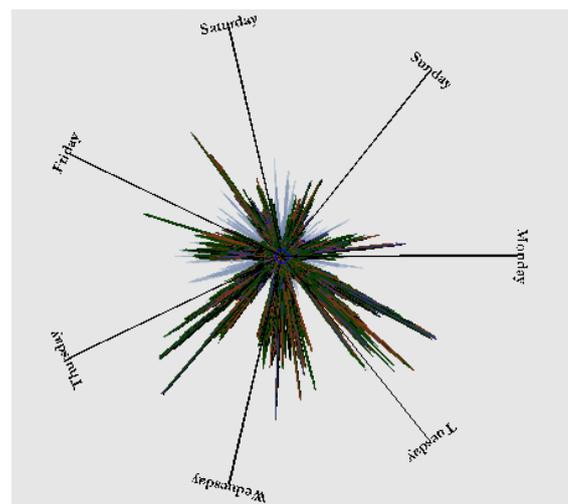


Fig. 6. Activity during the rest of the days of the week

C. Results of the application of the Semantic Tagcloud

The STC helps the people in charge of the course to analyze the interaction and content published during the course. In this case study have been used 5 academic years with an average of 18 forums, 123 discussion threads, 242 posts and 32105 forum accesses per academic course. This means an average of 3512 different tags per year. An example of the STC application for an academic year is shown in Fig 2.

Some of the terms shown are repeated along the five academic courses. Taking into account the different recorded information it is possible to assert that forums are used for:

- Publishing information. The forum is used to inform students about events related to the subject, such as: announcements, schedule changes, dissertation defense, list of assignment submissions and so on. Following words must be highlighted in this sense: *Noticia, Aviso, Defensa, Convoca, Examen, Hora*, etc.; which in English means News, Notice, Defence, Summons, Review, Time, respectively.
- Modelling Discussions. The forums are used to discuss issues and tools related to Software Engineering subject, tags such as: *Discusión, Entidad, Relación, Datos, Diagrama, Atributo, Modelo, SET*, etc.; which in English means Discussion, Entity, Relation, Data, Diagram, Attribute, Model, SET (that is the acronym of a used CASE tool), respectively. Also here some issues of the workshop statement must be considered, such as: *Id, Gestión, Sistema*, etc.; which in English means Identification, Management, System, respectively.
- Doubts Resolution. This is a very common use of the forum. Some tags related to this are: *semana, problema, prácticas, duda, forma, despacho, entrega, solución*, etc.; which in English means week, problems, practices, doubt, form, Web, Office, Delivery Solution, respectively.

D. Results of SNG application

The analysis of the participation is very important in this kind of subject. During the course a higher participation can be rewarded with a higher grade, however if students' interaction is low their grade will not be penalized.

This means that is important to see if there are students that are participating more in forums and accessing more to resources. To do so the SNG is very useful. For instance Fig 7 shows which are the most active users (the biggest) and the most important forums (the nearest to the users cloud and biggest) and the posts more read and resources more used (also the biggest). Furthermore the user of the SNG can hide and show connections between items. In this case it would be interesting to see the interaction between users Fig 3.

Another interesting functionality that the SNG provides is to know who are the students that use more the published resources. In this case a resource can be selected, drag to a corner of the screen and those users with any interaction with it will move toward the resource, the nearer a user is to the resource a higher interaction he/she has with it (Fig. 8).

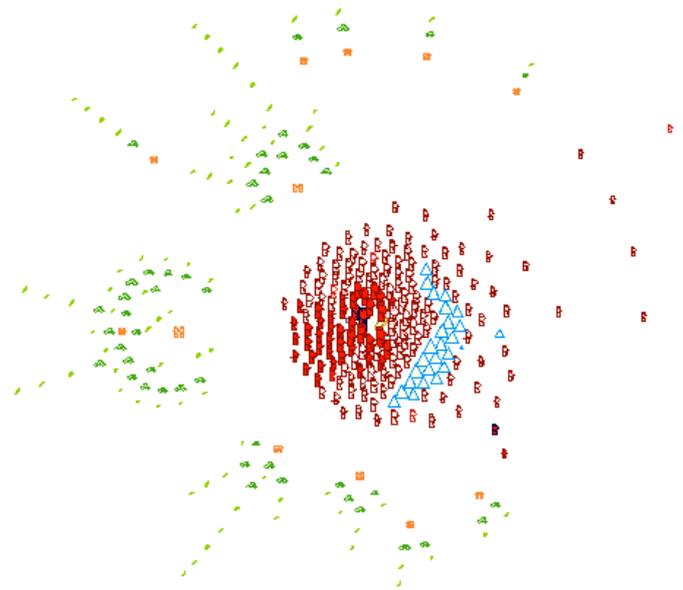


Fig. 7. Social Graph with the distribution of resources, users, forums and posts

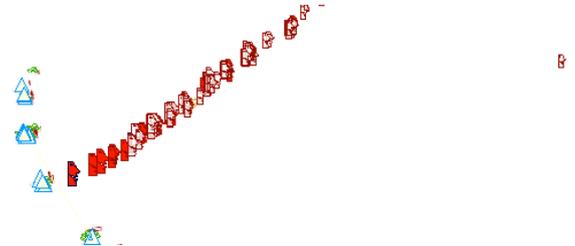


Fig. 8. Representation of users interaction with resources.

It would be possible to color the students taking into account their grades, but this has no sense for Software Engineering subject, because grades are not included into Moodle. To know if the interaction with resources can be related to higher grades it was necessary to select manually each student, get his/her data and check their grade also manually in the gradebook. With this strategy it was possible to see that the 72,8% of the students with a higher interaction have also better grades. To achieve this conclusion 20 of nearer students to the most relevant resources in the social graph were considered. If their grade was between 7-10/10 the relationship was confirmed. The percentage for this relationship in each academic year is shown in table 1.

E. Results of the application of the Parallel coordinates representation

In order to check these results and specially to facilitate the estimation of relationships between different dimensions Parallel coordinates can be applied.

TABLE I. PERCENTAGE OF STUDENTS WITH HIGH GRADES AND HIGH INTERACTION WITH RESOURCES

Ac. Year	08/09	09/10	10/11	11/12	12/13
Percentage	72%	67%	78%	65%	82%

However, in this case one dimension cannot be used, the grades. This is because (as commented above) Software Engineering subject do not use Moodle gradebook to manage students grades, so they are not in the LMS and they cannot be compared. Fig. 4 shows that all lines converge in the same point at the bottom of the last column. This is because such column represents the grades and the students' grades have not been introduced in the platform so their default grade for a course is 0. In case that grades were stored in the LMS this tool could help the responsible of the course to detect how participation impacts in grades, which are the most common interactions, which of them influences more, etc.

VI. CONCLUSIONS

The analysis of the information and interaction of students and teachers during the learning processes may facilitate making decision. However this is not an easy task taking into account the quantity of information stored in an LMS. In order to make this possible analytic techniques are applied and they are supported by systems such as VeLA. The case study has shown that VeLA makes possible to see:

- How the participation evolves over the time. It is possible to obtain knowledge about what are the most important moments in the academic course and introduce new staff or tools to support teaching and learning. Also, VeLA STT provides information about other moments with less participation, which can be enhanced by including new tools or contents, in order to have a similar participation throughout the overall course.
- What contents are more relevant for students. By using the STC it is possible to analyse the forum contents and discover the most important concepts, threads and forums. With this information, it is feasible making decisions about the inclusion of new topics, tools, activities and so on, when these elements may help students to grasp or empower difficult concepts.
- How students and teachers participates in the learning context. By using the SNG it is possible to see which are the users with more participation, with whom they are interacting, how the interact with resources. With this information it is possible to define strategies in order to improve learners' and teachers participation.
- The relationships between different dimensions of the interaction stored in the LMS. By using the Parallel coordinates tool it is possible to check how some kind of activities have a positive or negative impact in grades, so it would be possible to promote or avoid those activities.

In addition these tools can be combined with LA techniques. For instance by the inclusion of categories of interaction as de defined by Agudo-Peregrina et. al [18]. These categories have been applied to VeLA in an experiment [51] in other to check the conclusions posed by the previous research.

As future research lines, it would be desirable: 1) to apply these tools to other subjects data and in other contexts; 2) to complete the existing data with the students grades in order to

look for relations between the different type of interactions and the grades; 3) define new representation models.

In order to conclude we can say that by applying visual analytics systems and visualization models it is possible to see information about how students and teachers carry out their teaching/learning activities and make decision from it.

ACKNOWLEDGEMENT

This work is partially supported by Spanish Science and Innovation Ministry through the oiPLE project (ref. TIN2010-21695-C02) and also by the Regional Council of Education of Junta de Castilla y León through the projects GR47 and MPLE (ref. SA294A12-2).

REFERENCES

- [1] M. P. Prendes, "Plataformas de campus virtuales de Software Libre: Análisis compartido de la situación actual de las Universidades Españolas," Informe del proyecto EA-2008-0257 de la Secretaría de Estado de Universidades e Investigación 2009.
- [2] P. Avgeriou, A. Papasalouros, S. Retalis, and M. Skordalakis, "Towards a Pattern Language for Learning Management Systems," *Educational Technology & Society*, vol. 6, pp. 11-24, 2003.
- [3] S. Buckingham-Shum and R. Ferguson, "Social Learning Analytics," *Educational Technology & Society*, vol. 15, pp. 3-26, 2012.
- [4] M. Á. Conde and Á. Hernández-García, "A promised land for educational decision-making?: present and future of learning analytics," presented at the Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality, Salamanca, Spain, 2013.
- [5] D. Laney, "3D Data Management: Controlling Data Volume, Velocity, and Variety," META Group February 2001.
- [6] R. Ferguson, "The State Of Learning Analytics in 2012: A Review and Future Challenges," The Open University, UK 2012.
- [7] P. Goldstein, "Academic Analytics: The Uses of Management Information and Technology in Higher Education," *EDUCAUSE*, vol. 8, 2005 2005.
- [8] P. D. Long and G. Siemens, "Penetrating the Fog: Analytics in Learning and Education (EDUCAUSE Review) | EDUCAUSE.edu," *EDUCAUSE Review*, vol. 46, 2011/10/September 2011.
- [9] T. Barnes, M. Desmarais, C. Romero, and S. Ventura, "Educational Data Mining 2009," in *Proceedings of the 2nd International Conference on Educational Data Mining*, Córdoba, Spain, 2009, pp. 1-3.
- [10] A. van Barneveld, K. E. Arnold, and J. P. Campbell, "Analytics in Higher Education: Establishing a Common Language," *EDUCAUSE 1*, 2012 2012.
- [11] M. Brown, "Learning Analytics: Moving from Concept to Practice.," in *ELI Briefs*, 2012.
- [12] J. J. Thomas and K. A. Cook, *Illuminating the Path: The Research and Development Agenda for Visual Analytics*: IEEE Computer Society Press, 2005.
- [13] D. A. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann, *Mastering the Information Age - Solving Problems with Visual Analytics*: Eurographics Association, 2010.
- [14] L. Johnson, S. Adams Becker, V. Estrada, and A. Freeman, *NMC Horizon Report: 2014 Higher Education Edition*. Austin, Texas: The New Media Consortium, 2014.
- [15] L. Derick, P. Abelardo, V. Luis de la Fuente, S. David, C. nchez de, and K. Carlos Delgado, "GLASS: a learning analytics visualization tool," presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada, 2012.
- [16] D. Amo, M. J. Casany, and M. Alier, "Google Analytics for Time Behavior Measurement in moodle," presented at the Sistemas y Tecnologías de la Información. Actas de la 9ª Conferencia Ibérica de Sistemas y Tecnologías de la Información, Barcelona, Spain, 2014.

- [17] R. Mazza and V. Dimitrova, "CourseVis: A graphical student monitoring tool for supporting instructors in web-based distance courses," *International Journal of Human-Computer Studies*, vol. 65, pp. 125-139, 2// 2007.
- [18] Á. F. Agudo-Peregrina, S. Iglesias-Pradas, M. Á. Conde-González, and Á. Hernández-García, "Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning," *Computers in Human Behavior*, vol. 31, pp. 542-550, 2// 2014.
- [19] E. A. Kimberly and D. P. Matthew, "Course signals at Purdue: using learning analytics to increase student success," presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada, 2012.
- [20] J. A. Ruipérez-Valiente, P. J. Muñoz-Merino, and C. Delgado-Kloos, "An architecture for extending the learning analytics support in the Khan Academy framework," presented at the Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality, Salamanca, Spain, 2013.
- [21] A. Kerly, R. Ellis, and S. Bull, "CALMsystem: A Conversational Agent for Learner Modelling," in *Applications and Innovations in Intelligent Systems XV*, R. Ellis, T. Allen, and M. Petridis, Eds., ed: Springer London, 2008, pp. 89-102.
- [22] L. Ali, M. Hatala, D. Gašević, and J. Jovanović, "A qualitative evaluation of evolution of a learning analytics tool," *Computers & Education*, vol. 58, pp. 470-489, 1// 2012.
- [23] A. Fortenbacher, L. Beuster, M. Elkina, L. Kappe, A. Merceron, A. Pursian, et al., "LeMo: A learning analytics application focussing on user path analysis and interactive visualization," in *Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), 2013 IEEE 7th International Conference on*, 2013, pp. 748-753.
- [24] S. Dawson, A. Bakharia, and A. Heathcote, "SNAPP: Realising the affordances of real-time SNA within networked learning environments," in *Seventh International Conference on Networked Learning 2010*, 2010, pp. 125-134.
- [25] S. Jose Luis, G. Sten, V. Katrien, and D. Erik, "Goal-oriented visualizations of activity tracking: a case study with engineering students," presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada, 2012.
- [26] S. Govaerts, K. Verbert, D. Dahrendorf, C. Ullrich, M. Schmidt, M. Werkle, et al., "Towards Responsive Open Learning Environments: The ROLE Interoperability Framework," in *Towards Ubiquitous Learning*, vol. 6964, C. Kloos, D. Gillet, R. Crespo Garcia, F. Wild, and M. Wolpers, Eds., ed: Springer Berlin Heidelberg, 2011, pp. 125-138.
- [27] D. A. Gómez-Aguilar, F. J. García-Peñalvo, and R. Therón, "Analítica Visual en eLearning," *El Profesional de la Información*, vol. 23, pp. 236-245, 2014.
- [28] R. Mazza and C. Milani, "GISMO: a Graphical Interactive Student Monitoring Tool for Course Management Systems," presented at the TEL'04 Technology Enhanced Learning'04 International Conference, Milan, Italy, 2004.
- [29] V. Podgorelec and S. Kuhar, "Taking advantage of education data: Advanced data analysis and reporting in virtual learning environments," *Electronics and Electrical Engineering*, vol. 114, pp. 111-116, 2011.
- [30] H. Zhang and K. Almeroth, "Moodog: Tracking Student Activity in Online Course Management Systems," *Journal of Interactive Learning Research*, vol. 21, pp. 407-429, 2010.
- [31] Dell-Cloud-Services. (08/07/2014). *Collaboration, Analytics, and the LMS: A Conversation with Stephen Downes*.
- [32] M. A. Chatti, A. L. Dyckhoff, U. Schroeder, and H. Thüs, "A reference model for learning analytics," *Int. J. Technol. Enhanc. Learn.*, vol. 4, pp. 318-331, 2012.
- [33] C. Doug, "The learning analytics cycle: closing the loop effectively," presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada, 2012.
- [34] W. Greller and H. Drachler, "Translating Learning into Numbers: A Generic Framework for Learning Analytics," *Educational Technology & Society*, vol. 15, pp. 42-57, 2012.
- [35] S. Ben, "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations," presented at the Proceedings of the 1996 IEEE Symposium on Visual Languages, 1996.
- [36] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon, "Visual Analytics: Definition, Process, and Challenges," in *Information Visualization*, vol. 4950, A. Kerren, J. Stasko, J.-D. Fekete, and C. North, Eds., ed: Springer Berlin Heidelberg, 2008, pp. 154-175.
- [37] M. J. Casany, M. Alier, M. Á. Conde, and F. J. García, "SOA Initiatives for eLearning: A Moodle Case," presented at the 23rd International Conference on Advanced Information Networking and Applications, AINA, Bradford, United Kingdom, May 26-29, 2009, 2009.
- [38] M. Á. Conde, A. Pozo, and F. J. García-Peñalvo, "e-Learning Services in Moodle 2.0," *CEPIS Upgrade.*, vol. 12, pp. 43-50, 2011.
- [39] M. Á. Conde, "Personalización del aprendizaje: Framework de servicios para la integración de aplicaciones online en los sistemas de gestión del aprendizaje," Doctorado en Informática y Automática Tesis Doctoral, Departamento de Informática y Automática, Universidad de Salamanca, Salamanca, 2012.
- [40] D. A. Gómez, R. T. Sánchez, and F. J. García, "Semantic Spiral Timeline as a Support for eLearning," *Journal of Universal Computer Science*, vol. 15, pp. 1526-1545, 2008.
- [41] D. Modjeska, "Navigation in Electronic Worlds: Research Review for Depth Oral Exam David Modjeska Department of Computer Science.," 1997.
- [42] D. A. Gómez, M. Á. Conde, R. Therón, and F. J. García-Peñalvo, "Reveling the evolution of semantic content through visual analysis," presented at the 11th IEEE International Conference on Advanced Learning Technologies (ICALT 2011), Athens, Georgia, USA, 2011.
- [43] Y. Hassan-Montero and V. Herrero-Solana, "Improving Tag-Clouds as Visual Information Retrieval Interfaces," presented at the I International Conference on Multidisciplinary Information Sciences and Technologies, Mérida, Spain, 2006.
- [44] L. Bongshin, N. H. Riche, A. K. Karlson, and S. Carpendale, "SparkClouds: Visualizing Trends in Tag Clouds," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, pp. 1182-1189, 2010.
- [45] D. Modjeska, "Navigation in Electronic Worlds: Research Review for Depth Oral Exam David Modjeska Department of Computer Science," ed, 1997.
- [46] L. Lipponen, M. Rahikainen, J. Lallimo, and K. Hakkarainen, "Patterns of participation and discourse in elementary students' computer-supported collaborative learning," *Learning and Instruction*, vol. 13, pp. 487-509, 10// 2003.
- [47] D. A. Gómez, R. Therón, and F. J. García, "Understanding educational relationships in Moodle with ViMoodle," presented at the The Eighth IEEE International Conference on Advanced Learning Technologies, 2008. ICALT 08, Santander, Cantabria, Spain, 2008.
- [48] L. S. Bershtein and A. V. Bozhenuk, "Fuzzy coloring for fuzzy graphs," presented at the The 10th IEEE International Conference on Fuzzy Systems Melbourne, Vic. , Australia, 2001.
- [49] A. F. Agudo-Peregrina, A. Hernandez-Garcia, and S. Iglesias-Pradas, "Predicting academic performance with learning analytics in virtual learning environments: A comparative study of three interaction classifications," in *Computers in Education (SIEE), 2012 International Symposium on*, 2012, pp. 1-6.
- [50] I. Alfred and D. Bernard, "Parallel coordinates: a tool for visualizing multi-dimensional geometry," presented at the Proceedings of the 1st conference on Visualization '90, San Francisco, California, 1990.
- [51] G.-A. Diego Alonso, F. J. García-Peñalvo, and R. Theron, "Tap into visual analysis of the customization of grouping of activities in eLearning," presented at the Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality, Salamanca, Spain, 2013.