

# Generation of Graphical Representations of Student Tracking Data in Course Management Systems

Riccardo Mazza  
Faculty of Communication Sciences  
University of Lugano, Via Buffi 13  
6900 Lugano - CH  
riccardo.mazza@lu.unisi.ch

Vania Dimitrova  
School of Computing  
University of Leeds, LS2 9NA, Leeds - UK  
vania@comp.leeds.ac.uk

## Abstract

*An approach of employing Information Visualisation to develop systems that facilitate instructors in web-based distance learning is presented here. The paper describes a tool, called CourseVis, that uses multidimensional student tracking data collected by CMS and generates graphical representations that can be used by instructors to gain an understanding of what is happening in distance learning classes. The work followed a systematic approach that started from collecting the instructors' needs, produced some appropriate graphical representations of student tracking data, and evaluated the effectiveness, efficiency and usefulness of the proposed representations. The evaluation has shown that with CourseVis the instructors can identify tendencies in their classes, quickly discover individuals that need special attention, and are able to provide better support to their students.*

## 1 Introduction

Web-based learning environments, often called Course Management Systems (CMS), are widely used nowadays in universities and companies around the world. These tools take advantage of computer mediated communication to create virtual classrooms where learners and instructors share a common learning space. With CMS, the instructors can produce and distribute content material, prepare assignments, engage in discussions with the students, and manage distance classes. The students have unlimited access to the learning material from any location and can submit assignments and carry out quizzes at a convenient for them time. The students often need help from the instructors to understand unclear concepts, and benefit from feedback on their progress. CMS have now become so popular that many universities around the world have adopted them as the main

platform for on-line learning. Some well-known commercial tools are WebCT and BlackBoard. Free, open source solutions, such as Moodle, are being often employed as alternatives to the more expensive commercial CMS.

Despite the fact that CMS have reached a good level of maturity and reliability, they still fail to provide effective support to the instructors to perform basic student monitoring activities. Indeed, educational research reports a number of problems with using CMS in distance learning, such as high student drop-out rates, students' feeling of isolation, lack of motivation, and difficulties with finding support from tutors and peers [16, 3, 15, 17, 4]. Some authors suggest that the regular monitoring of the students' activities can reduce on-line distance learning problems [2, 14, 3]. This includes monitoring of a learner's progress with the material and testing of the acquired knowledge and skills on a regular basis. Furthermore, the instructors should be provided with appropriate means to monitor the students' activities and to diagnose particular situations that might require the instructor's intervention to prevent or overcome potential problems.

CMS store in their internal databases large log files with the students' activities in the course, and have a built-in student monitoring feature that enables the instructors to view some statistical data, such as the number of accesses made by the students to each resource, the history of pages visited, and the number of hits for every day of the course. However, this information is usually provided in a tabular format, often incomprehensible, with a poor logical organization, and is difficult to interpret. As a result, web log data is very rarely used by distance learning instructors.

In this work, we propose an approach to graphically represent student tracking data generated by CMS. We have implemented a tool, called CourseVis, that uses multidimensional student tracking data collected by CMS and generates appropriate graphical representations that can be used by distance learning instructors to gain an understanding of

their on-line students.

The paper is organized as follows. Firstly, we will justify our research and discuss related approaches. Section 3, which is the core of the paper, will describe the CourseVis system and some of the graphical representations generated. Findings from the CourseVis evaluation with instructors will be outlined in Section 4. Finally, in the conclusions, we will summarise and sketch out some ideas for future work.

## 2 Visualizing student tracking data

This work is an attempt to apply some Information Visualisation concepts and ideas to deal with a vast amount of student tracking data collected by CMS. Visualisation techniques have already been used to illustrate student data, but although a great number of successful visualisation techniques and systems have been developed, the application of this research into distance learning has not been explored in a systematic way. There is a small number of recent projects in computer-based learning that use visualisation techniques to present data to users. For example, visualization techniques have been used to present social aspects in computer-supported collaborative learning [15], community relationships in peer-to-peer systems [1], and conversations in online groups [6]. These projects commonly explore only one graphical representation to solely present data in a pictorial way and provide no operations to manipulate the graphics. The application of Information Visualisation is very limited: graphical representations are merely used as tools for communicating information to others. The advantages of Information Visualisation for exploring data, discovering tendencies, confirming or rebutting hypotheses have not been exploited systematically in computer-based learning, in general, and in distance learning, specifically.

Some effort has been dedicated to explore forms of visualising cognitive aspects of students with open student models [18, 7, 5]. These works use pictorial representations to externalise a student model, which is a data structure built by the system based on some Artificial Intelligence inference. In general, maintaining dynamic student models is a technically challenging task. Once such models are built, the visualisation techniques used to show the models to the students are fairly basic. Moreover, open student models explore graphical representations to show only one aspect of the students' data, namely the students' cognitive state.

With CourseVis, we want to go further to provide multiple representations of the data in order to show different perspectives that allow the instructors to examine cognitive, social, and behavioural aspects of distance students. By the means of visualisation techniques, CourseVis will transform the multidimensional student tracking data into graphical representations to help instructors build mental models

of what is happening in their classes.

Our research followed a systematic approach that included requirements gathering, design, implementation, and evaluation. To find out what information about distance students the instructors need when they run courses with CMS, as well as to identify possible ways to help instructors to acquire this information, we conducted a survey with instructors involved in distance learning. The results of the survey, which are presented elsewhere (see [12]), drew the requirements for the design of CourseVis. Based on these requirements, we designed the graphical tools and implemented the system, as described in the next session. The educational rationale for the system was discussed in [11]. Finally, an empirical evaluation was conducted with potential users of the system taking into account the tasks that instructors would perform in their practice. The evaluation focused on the effectiveness, efficiency, and usefulness of CourseVis and is briefly outlined in this paper, detailed description is given elsewhere, see [10].

## 3 The CourseVis architecture

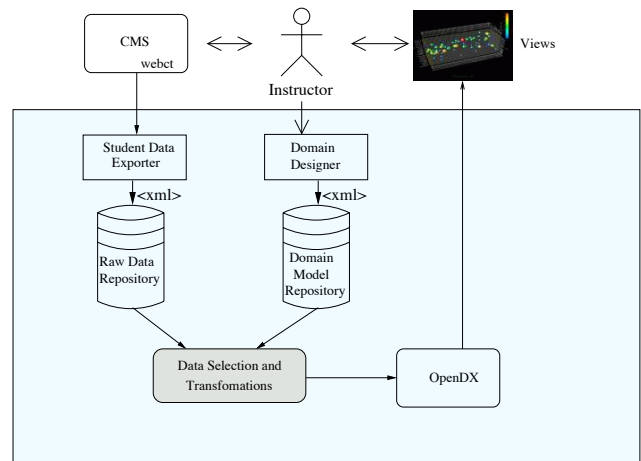


FIGURE 1: The CourseVis system architecture

CourseVis has been designed as a tool that extends an existing CMS (we used WebCT in our current implementation) by adding new functionality in terms of graphical representations produced with OpenDX [13] - a powerful, open source, generic visualisation tool. The architecture of CourseVis is presented in Figure 1.

CourseVis includes a *Student Data Exporter* that collects and transforms the student tracking data provided from CMS in a well-defined XML format. This XML data is then stored in the *Raw Data Repository*. A specific implementation of the Student Data Exporter for the WebCT course management system was built.

Some visualisations require that students are related to

entities of the course domain (called *concepts*). For instance, the results of quizzes could be used to judge the students' level of mastering of domain concepts. In order to represent the level of knowledge achieved by each student on a specific concept, CourseVis uses a pre-defined *domain model*. Such a model enables the linking of content material and evaluation proofs (i.e. quizzes) with concepts of the course. A *Domain Designer* module facilitates the instructors to create the domain model of the course. The domain model defines a list of domain concepts and specifies a set of associated concepts for each page of the course and for every question in quizzes. A *Domain Model Repository* stores the XML data produced by the Domain Designer.

To represent the data in a visual format, CourseVis performs some computations and transforms the data into visual structures and representations. The *Data Selection and Transformations* module is in charge of:

- extracting from the repositories only data relevant to a particular representation to be generated;
- manipulating the data extracted from the repository and transforming it in a format to be processed by the OpenDX visualisation tool.

CourseVis generates several graphical representations, described in detail in [10], some of which are illustrated below. The data used is from an on-line course in Java Programming ran by the first author at the Department of Informatics and Electronics of the University of Applied Sciences of Southern Switzerland.

The preliminary study that informed the design of CourseVis [12] identified what information the instructors would need in their on-line teaching activities. This was classified in social, cognitive and behavioural aspects, the visualisations for which are described next.

### 3.1 Visualising social aspects of learners: Discussion Plot

The instructors indicated that information about social activities of students based on interactions using the discussion boards in CMS was important for managing on-line classes. The discussion board is a tool which allows students to read and post messages. Each message has a *sender*, *date*, and *topic*. A set of posts on the discussions, composed of an initial post about a topic and all responses to it is called a *thread*. We call the person who started the initial posting in a thread *originator*. The length of a thread (i.e. the number of messages written on a specific subject) is called *follow-up* and is sometimes useful for monitoring discussions. These variables (originator, date, topic, and follow-up) are mapped onto a 3-D scatterplot illustrated in

Figure 2 <sup>1</sup>.

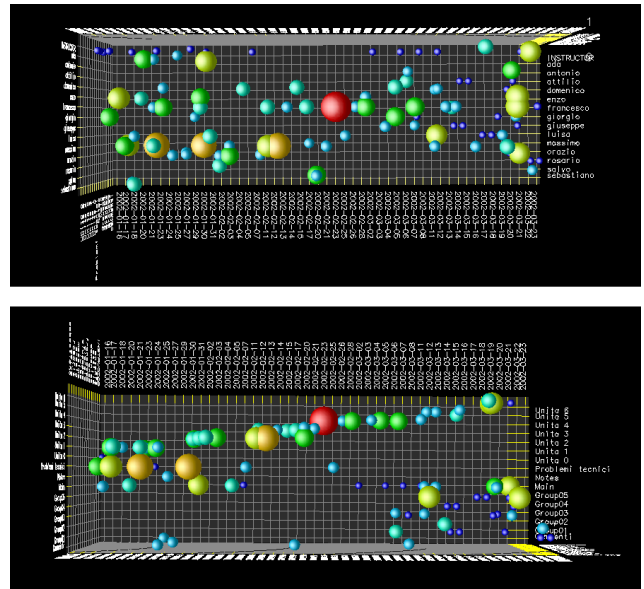


FIGURE 2: Discussion Plot example: The top scatterplot presents a graphical representation of discussion threads showing the students who have initiated the threads; at the bottom, the same scatterplot is rotated to show discussions related to the course topics.

In the image, the size of the sphere represents the number of follow-ups in a discussion. The instructor may use operations, such as zooming and panning, to analyse the students' engagement in discussions. For instance, Figure 2 shows two rotations that enable the analysis of different relations by examining the same data set. The figure allows the instructor to identify certain social characteristics: students who dominated in opening new threads (*Francesco* and *Massimo* in the examples), students who have not initiated any discussions, and the size of discussions on different topics of the course.

### 3.2 Visualising cognitive aspects of learners: Cognitive Matrix

CMS provide the instructors with tools, such as quizzes and assignments, to measure the level of comprehension achieved by the students on the domain concepts. The survey with instructors indicated that they needed information about the overall student performance, as well as the students' performance on specific domain concepts.

One of the visualizations that CourseVis provides is the Cognitive Matrix (see Figure 3). It uses the students' performance on quizzes to assess the students' understanding of

<sup>1</sup>In this and the following images the student names are not real due to data protection considerations

the subject. The student names are mapped onto the x-axis and the concepts of the course are mapped onto the y-axis. The performance values are mapped onto the colour of the square corresponding to a student and a concept. If a student has not attempted the questions related to a concept, the corresponding cell is blank.

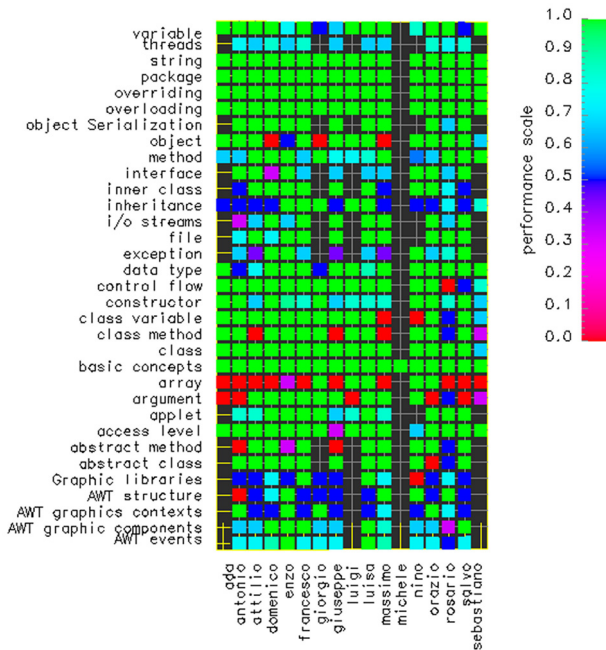


FIGURE 3: The Cognitive Matrix visualises the students' performance on quizzes related to domain concepts.

The matrix allows micro and macro analysis of the students' performance. The instructor may use it to detect problematic topics or struggling students by comparing columns and rows, or can analyse the performance of a particular student on a specific topic. In CourseVis, several graphical representations can be combined for a more comprehensive analysis. For instance, the Discussion Plot shows that the student *Massimo* has opened long discussion threads, while the Cognitive Matrix shows that his performance on quizzes was not that good.

### 3.3 Visualising the learners' access to the course: Student Accesses Plot

The survey with instructors indicated that the monitoring of the students' attendance was important for the effective managing of distance courses. The *Student Accesses Plot* is generated in CourseVis to enable the monitoring of the students' access to the course. Every time a *student* accesses the course, the CMS registers in its internal data structures the *date* and the *duration* of the access. This data is used

to create visual representations that summarise the overall students' attendance to the course. The instructors may be interested also to know how many times the course has been accessed by all students on a particular day *D* of the course.

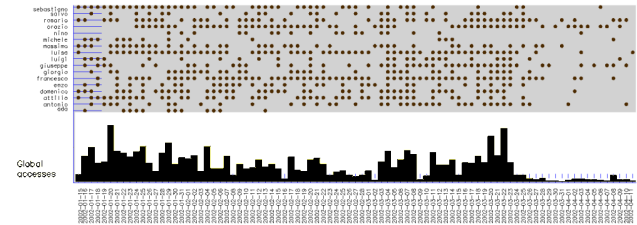


FIGURE 4: The Student Accesses plot, represents the students' attendance to the on-line course.

The student access plot in CourseVis is a composition of a scatterplot and a histogram, see Figure 4. The scatterplot represents bivariate data of students and dates of the course, where a corresponding bullet represents at least one access to the course made by the student on the corresponding date. The histogram represents the number of pages of the course accessed by all students on a particular date.

The instructor has an overview of the global student access to the course with a clear identification of patterns and trends. For instance, there is a high concentration of access between the 4th and 23rd of March, while little access is made at the end of the course. It is also possible to have attendance information about a specific student, e.g. Michele made little visits to the course material, while Francesco was quite active throughout the course.

### 3.4 Visualising behavioural aspects of learners: Behaviour Graph

The survey indicated that representations of the students' behaviour, including whether the students read the materials, perform evaluation proofs, and participate in discussions, as well as the students' progress with the schedule of the course, will be helpful for the instructors.

This aspects are covered in the *Student Behaviour Graph*, illustrated in Figure 5. It represents information regarding a particular student, and takes advantage of the single-axis composition method [8] for presenting a large number of variables in a 2D metric space. With a common x-axis mapping the dates of the course, a number of variables are represented:

- *Accesses to content pages by topics.* A square is used to mark that a student has accessed a content page with material linked to a particular concept. The instructor can see the concepts studied by the student and can find out whether he/she has missed any concepts.

# Summary of student's behaviours

from 2002-01-15 to 2002-04-11  
Student: "francesco"

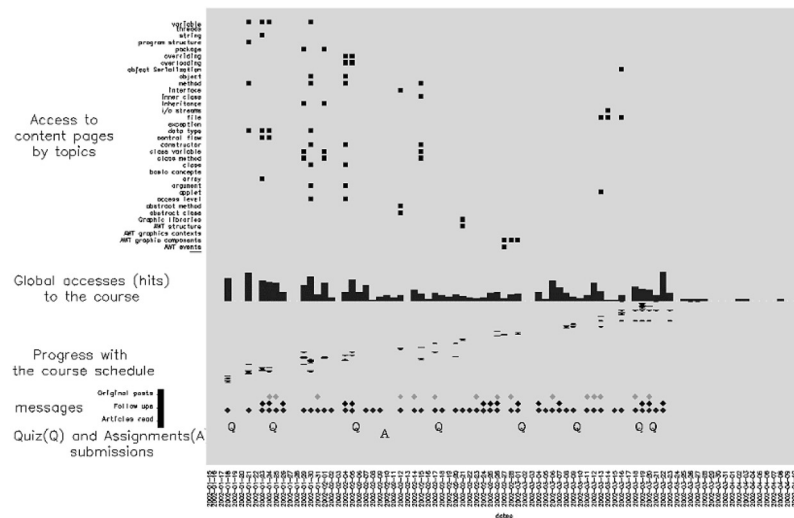


FIGURE 5: The Behaviour Graph, represents the student's behaviour.

- *Global accesses to the course.* The number of accesses (including any part of the course, not only content materials) on a single day is shown on the histogram.
- *Progress with the course schedule.* We assume that the course has a linear structure, i.e. it has a syllabus with a list of content pages having a schedule organised in a linear sequence (the table of contents). We put in the y-axis the page ordering according to this sequence, a mark then represents a page accessed by the student on a day, where the y-location of the mark is proportional to the sequence of the pages in the schedule (first page on bottom, last page on top). The instructor may have an overview of the progress made by the student with the schedule during the time.
- *Messages.* Some information about the participation in discussions has been already considered in the discussion plot. However, some further information on discussion attendance could be given. In particular, information about the posting of a new message, the reply to a message, or simply the reading of a message, is denoted in the representation with a diamond coloured in a different colour (e.g. red, green and blue).
- *Quiz and assignment submission.* The submission of a quiz or an assignment is reported in the graphic on the date it was submitted with a Q or A.

## 4 Evaluation with instructors

An empirical evaluation of CourseVis was conducted [10] focusing on its effectiveness (can it help instructors gain an understanding of what is happening in distance classes), efficiency (can instructors infer required information quickly), and usefulness (to what extent the information provided is useful to the instructors). The evaluation involved instructors with experience in using CMS in distance learning, and took place in three stages. 1) A focus group was conducted with five instructors who were provided with CourseVis. The aim was to identify problems with the representations generated in CourseVis and assess their usefulness. 2) An experimental study was conducted with six instructors with experience in distance learning. The instructors were asked to imagine that they were running the course and had to find some information about the students in order to answer a set of pre-designed questions. Instructors were split in two groups: 3 used a standard CMS platform, and the other 3 used the same platform and tools, but integrated with CourseVis. 3) Finally, semi-structured interviews were conducted with the three participants who used CourseVis in the experimental study to discuss the usefulness of the graphical representations.

The evaluation of CourseVis showed that the representations helped the instructors to quickly and more accurately grasp information about social, cognitive, and behavioural aspects of the students. The provided information was regarded by the teacher as very useful for managing success-



ful distance courses. It was noted that the graphical representations provided in CourseVis may help instructors identify early, and even prevent, some of the problems with distance learning, e.g. students who do not communicate might feel isolated, a student not visiting the course material might be confused or be a potential drop-out, long discussion threads on a topic may highlight problems experienced by learners (e.g. technical aspects).

## 5 Conclusions

This paper presented a novel approach of using web log data generated by Course Management Systems to help instructors become aware of what is happening in distance learning classes. Specifically, techniques from information visualization have been used to graphically render complex, multidimensional student tracking data collected by CMS. A system, called CourseVis, illustrates the proposed approach. CourseVis has been developed as a generic tool for visualising student tracking data. It is required that the data from CMS are converted into an xml format, which is then imported into the CourseVis graphical render to produce various representations of the data. Some representations from CourseVis instantiation have been presented in the paper to illustrate the use of information visualization for rendering student tracking data. The validity of our approach was confirmed with the evaluative studies of CourseVis. The system was regarded as useful for helping the instructors form a mental model of their classes and offer an appropriate help to the students.

Our immediate research plans are to address some of the problems revealed in the CourseVis evaluation. These include providing flexible links between the graphical representations, ensuring a connection between graphics and corresponding data from CMS (e.g. a link from a discussion graph to postings in the discussion forum), enabling instructors to specify what data should be included in or excluded from graphics. Some of these issues are addressed in the successor of CourseVis - GISMO [9] a visualisation tool for student monitoring being developed within the EU funded EdukaLibre project. GISMO is integrated in the Moodle open source learning environment and is currently being evaluated in several studies with real users.

## References

- [1] H. Bretzke and J. Vassileva. Motivating cooperation on peer to peer networks. In *Proceedings of the 9th International Conference on User Modeling*, number 2702 in LNAI. Springer, 2003.
- [2] K. Cotton. Monitoring student learning in the classroom. School Improvement Research Series (SIRS). Northwest Regional Educational Laboratory, U.S. Department of Education. <http://www.nwrel.org/scpd/sirs/2/cu4.html>, 1988.
- [3] J. M. Galusha. Barriers to learning in distance education. <http://www.infrastructure.com/barriers.htm>, 1997.
- [4] N. Hara and R. Kling. Student distress in web-based distance education course. *Information, Communication & Society*, 4(3):557–579, 2000.
- [5] D. Hartley and A. Mitrovic. Supporting learning by opening the student model. In *Intelligent Tutoring Systems, 6th International Conference, ITS 2002, Biarritz, France and San Sebastian, Spain, June 2-7, 2002, Proceedings*, volume 2363 of *Lecture Notes in Computer Science*, pages 453–462. Springer, 2002.
- [6] D. Judith. A semantic approach to visualizing online conversations. *Communications of the ACM*, 45(4):45–49, April 2002.
- [7] J. Kay. The um toolkit for cooperative user modelling. *User modeling and user adapted interaction*, 4:149–196, 1995.
- [8] J. D. Mackinlay. Automating the design of graphical presentations of relational information. *ACM Transaction on graphics*, 5(2):110–141, 1986.
- [9] R. Mazza. Gismo: a graphical interactive student monitoring tool for course management systems. In *T.E.L. Technology Enhanced Learning Conference, Milan*, 2004.
- [10] R. Mazza. *Using Information Visualisation to Facilitate Instructors in Web-based Distance Learning*. PhD thesis, University of Lugano, 2004.
- [11] R. Mazza and V. Dimitrova. Coursevis: Externalising student information to facilitate instructors in distance learning. In U. Hoppe, F. Verdejo, and J. Kay, editors, *Proceedings of the International conference in Artificial Intelligence in Education. Sydney July 20-24, 2003. (AIED 2003)*, pages 279–286. IOS press, 2003.
- [12] R. Mazza and V. Dimitrova. Informing the design of a course data visualiser: an empirical study. In *Proceedings of the 5th International Conference on New Educational Environments. Lucerne 26-28 May. (ICNEE 2003)*, pages 215 – 220, 2003.
- [13] Opendx - open visualization data explorer, 2004. <http://www.opendx.org>.
- [14] L. C. Ragan. Good teaching is good teaching: an emerging set of guiding principles and practices for the design and development of distance education. *DE-OSNEWS, The American Center for the Study of Distance Education, Pennsylvania State University*, 8(12), 1998. <http://www.educause.edu/ir/library/html/cem9915.html>.
- [15] J. C. Rivera and M. L. Rice. A comparison of student outcomes & satisfaction between traditional & web based course offerings. *Online Journal of Distance Learning Administration. The State University of West Georgia.*, 5(3), 2002.
- [16] K. Smith-Gratto. Best practices and problems north carolina a & t state university. Report to the Distance Education Evaluation Task Force Distance Education. <http://qed.ncat.edu/ir&p/report.htm>, 1999.
- [17] D. Valentine. Distance learning: Promises, problems, and possibilities. *Online Journal of Distance Learning Administration*, 5(3), 2002. State University of West Georgia, Distance Education Center.
- [18] J. D. Zapata-Rivera and J. E. Greer. Visualizing and inspecting bayesian belief models. In *International Joint Conference on Artificial Intelligence IJCAI*, pages 47–49, 2001.