Goal-oriented visualizations of activity tracking: a case study with engineering students

Jose Luis Santos, Sten Govaerts, Katrien Verbert, Erik Duval
Departement Computerwetenschappen, K.U.Leuven
Celestijnenlaan 200A
B-3001 Leuven, Belgium
{joseluis.santos, sten.govaerts, katrien.verbert, erik.duval} at cs.kuleuven.be

ABSTRACT
Increasing motivation of students and helping them to reflect on their learning processes is an important driver for learning analytics research. This paper presents our research on the development of a dashboard that enables self-reflection on activities and comparison with peers. We describe evaluation results of four iterations of a design based research methodology that assess the usability, use and usefulness of different visualizations. Lessons learned from the different evaluations performed during each iteration are described. In addition, these evaluations illustrate that the dashboard is a useful tool for students. However, further research is needed to assess the impact on the learning process.

Categories and Subject Descriptors
H.5.2 [Information interfaces and presentation]: User interfaces; K.3.2 [Computers and Education]: Computer Science Education

General Terms
Design, Experimentation, Human Factors

Keywords
Learning analytics, Visualization, Reflection

1. INTRODUCTION
Increasing student motivation and assisting students with self-reflection on their learning processes is an important driver for learning analytics research. Student motivation can improve when students can define their own goals [30]. Visualizations of time spent and resource use can improve awareness and self-reflection [15]. Learning management systems (LMS) track most of the user interaction that can be used for learning analytics. However, many of the activities take place outside of the LMS, such as brainstorming or programming activities.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

LAK ’12 29 April - 2 May 2012, Vancouver, BC, Canada.
Copyright 2012 ACM 978-1-4503-1111-3/12/04 ...$10.00.

This paper presents the first results of a case study in a “problem solving and design” course for second year engineering students at the Katholieke Universiteit Leuven. In this course, the students have to develop software and go through the different phases of software development process, such as design, programming and reporting. To this end, they use tools such as LibreOffice\(^1\), the Eclipse IDE\(^2\) and Mozilla Firefox\(^3\). They have to share tasks and responsibilities between group members. Controlling the risks and evolution of such tasks is part of the assignment.

We developed a dashboard with visualizations of activity data. The overall goal of this dashboard is to enable students to reflect on their own activity and compare it with their peers. The time spent with different tools, websites and Eclipse IDE documents are tracked by RescueTime\(^4\) and the Rabbit Eclipse plug-in\(^5\). The collected information is displayed in a dashboard containing goal-oriented visualizations. In the visualizations, the students can filter by different criteria, such as course goals and dates. Such filters allow contextualization of the visualized data for the user. Linking the visualizations with the learning goals can help students and teachers to assess whether the goal has been achieved [12].

The dashboard is developed using the design-based research methodology, which relies on rapid prototyping, deployment of artifacts and observation in iterative cycles. The paper is organized as follows: in the next section, we present related work. Section 3 presents the research methodology. The four iterations of the design process are discussed in sections 4, 5, 6 and 7. Future work and conclusions are presented in Section 8 and 9.

2. RELATED WORK
Learning analytics considers the analysis of communication logs [33, 6], learning resources [25], learning management system logs and existing learning designs [21, 32], and the activity outside of the learning management systems [29, 9]. The result of this analysis can be used to improve the creation of predictive models [37, 13], recommendations [42] and reflection [15].

This paper focuses on activity outside learning management systems using existing tracking tools. Self-tracking

\(^1\)LibreOffice, http://www.libreoffice.org/
\(^2\)IDE Eclipse, http://eclipse.org
\(^3\)Mozilla Firefox, http://www.mozilla.org/
\(^4\)RescueTime, http://rescuetime.com/
tools can be used for capturing activities of students with different tools. The goal is to help students learn how they and other students are using their applications to achieve concrete goals.

Self-tracking is becoming popular in many domains, including Personal Informatics [20]. Applications in these domains help people understand their habits and behavior, through tracking and visualization, e.g., for sleep tracking and email communication patterns. Tracking of health data can motivate users with fitness exercises [31, 4] and enable early diagnosis of illness [39, 2, 1, 16]. Within companies, tracking and visualizations are used to analyze business and manufacturing processes [35], as well as productivity [3]. Behind these tools are communities where users can share experiences, publish their tracking data in social networks or compare the data with others. In a learning context, students and teachers are part of a community. These tools can play an important role to share and learn from their behavior with applications to achieve the goals of the course.

Khan Academy enables tutors to check progress of students [8]. A dashboard is used where a table provides a goal status overview per student. For every student, a timeline shows the distribution of achieved goals and a bar chart visualizes the time spent with different kinds of resources.

Other learning dashboards use pie charts to describe the access distribution of different user roles, simple plots to express time spent and tables to indicate the success rate for assignments [23]. In adaptive learning environments, dashboards contain box plots to compare grades and averages of users who have followed different paths [5]. In mashup learning environments, pie charts have been used to represent knowledge in different areas [24]. Tree visualizations are useful to express learning paths and to describe prerequisites. Each path can represent a knowledge area or subarea in a domain [26, 27]. In addition, there are models exploring ways to analyze electronic traces to create group models that can operate as mirrors which enable the individuals and teams to reflect on their progress through visualizations [41, 18].

The work presented in this paper focuses on tracking activity from different applications. Our dashboard uses different trackers that generate different kinds of data and applies different visualization techniques. The overall goal of this mashup of visualizations is to enable students to learn how they are using the tools and how much progress they make towards goals in comparison with peers.

3. RESEARCH METHODOLOGY

The design-based research methodology has been applied to conduct this research. This methodology relies on rapid prototyping to evaluate ideas in frequent short iteration cycles [43]. The approach enables to collect both qualitative and quantitative evaluation data during the whole software design process [28].

In the two first iterations, we developed a paper-based and a digital prototype. The evaluation of those iterations collected qualitative data from interviews and user observations of 15-30 minutes using the think-aloud protocol [19]. Six teachers and teaching assistants participated in the first iteration and 5 in the second iteration. Evaluations with these participants are useful to collect requirements and to identify potential usability issues with the interaction techniques.

The third and fourth iteration are conducted by a mixed research evaluation methodology with questionnaires and open ended questions. In these iterations, we conducted the evaluations with 36 and 10 students, respectively. These questionnaires focused on concrete aspects of the application and allowed statistical analysis of the evaluation data.

4. PAPER PROTOTYPE

Paper prototyping is an important first step in user interface design to get quick feedback [28] and minimize costs in the software design process [11].

4.1 Design and implementation

User activities and their visualization need to be related to learning goals in order for teachers or students to be able to reflect and make decisions. Linking these visualizations to the intended goals allows to assess whether these goals have been achieved [12].

The design of the paper prototype focuses on the above-mentioned use cases: (a) Students can reflect on different visualizations contextualized by the learning goals of the course and (b) enable social support through communication between students and teachers.

The first use case is addressed by using visualizations of user activities. More specifically, we visualize the behavior of students with different tools they are using for course activities (e.g., Eclipse IDE for programming and Microsoft Word for writing) to gain insight into what students have done to achieve a goal. To this end, the students and status of the goals are visualized as a table (visualization 1 Figure 1), as such visualization is one of the simplest ways to get an overview of the course [36]. Students are listed in the first column and the rest of columns represent the goals of the course and their status. A timeline [17] with bubbles (visualization 3) represents the number goals over time. Visualization 5 is a timeline that shows the number of events or time (depending of the tracked source of information) per tool that the user has used along the time to achieve the goal. Finally, visualization 5, 7 and 8 display time spent or number of events per weekday, actions and different items.

All visualizations together in the dashboard enable filtering information from a generic perspective (table with goals and students) to a more detailed description (type of documents and activity used), following the visual information seeking mantra [38]. This prototype enables users to interact with the visualizations, i.e., if a user clicks on the Monday bar (visualization 6), then visualization 7 and 8 displays related information to the clicked day.

All visualizations together in the dashboard enable filtering information from a generic perspective (table with goals and students) to a more detailed description (type of documents and activity used), following the visual information seeking mantra [38]. This prototype enables users to interact with the visualizations, i.e., if a user clicks on the Monday bar (visualization 6), then visualization 7 and 8 displays related information to the clicked day.

The second use case is addressed by providing chat functionalities for communication between teachers and students (number 2 and 4 in Figure 1). To enable social support, communication between users and sharing experiences can help users to achieve their goals. We provide two widgets intended for this purpose. One widget shows publicly the message and the other is reserved for private communication. The communication is always related to a specific goal contextualizing the scope of the conversation.

4.2 Evaluation

4.2.1 Demographics and Evaluation Setup

Users were interviewed and observed during 15-30 minutes. They had to perform different predefined tasks such
Figure 1: Paper prototype

The paper prototype was evaluated with six people (1 female and 5 males computer science teachers and assistants). Three participants were between 25-30 years old and two participants between 40-50.

4.2.2 Evaluation results and discussion

In this subsection, we introduce first the more remarkable problems and suggestions of the users, and finally the proposed solutions.

Three issues were highlighted with visualization 1. First, the headers in the table are the titles of the goals. The size of the table increases proportionally to the number of goals. If the number of goals is high, the user will not be able to get easily an overview due to the size of the table. Although the user can filter the goals on the table restricting the period of time, it requires additional steps for the user and affects the usability of the application. Second, the filtering feature is defined by drop-down lists for the day, month and year and requires too much clicks. Third, the table is showing a pop-up with static information when the mouse hovers over a cell. Pop-ups showing always the same information were identified as redundant by the users. In addition, users requested more sorting options for the table.

There were several problems to understand visualization 3. The visualization shows redundant information compared with the table. The table also includes goals and users can filter by time, so they can obtain the same information with this visualization. Users expected some additional information that they did find in the table. Although the user can filter the goals on the table restricting the period of time, it requires additional steps for the user and affects the usability of the application. Second, the filtering feature is defined by drop-down lists for the day, month and year and requires too much clicks. Third, the table is showing a pop-up with static information when the mouse hovers over a cell. Pop-ups showing always the same information were identified as redundant by the users. In addition, users requested more sorting options for the table.

Regarding visualization 3, we propose to replace it by a motion chart visualization. Such a visualization allows to show the evolution of the user activity and to compare it with the average of a group over time. In this way, we provide additional information as requested by the users.

5. DIGITAL PROTOTYPE

This iteration focuses on addressing usability issues de-
tected in the previous iteration.

5.1 Design and implementation

Dashboard personalization is provided by using widget technology. Such technology enables easy addition and deletion of widgets with different visualizations (see figure 2 to see the different technologies). We use the OpenSocial specification\(^6\) to enable deployment in OpenSocial compliant widget containers, e.g. iGoogle\(^7\) or Apache Shindig\(^8\).

OpenSocial enables inter-widget communication via OpenApp\(^9\) that allows to send information from one widget to another. User interactions with one widget are broadcasted to other widgets. These widgets can then also act upon these events, i.e. to filter data. Furthermore, iGoogle supports the concept of spaces. These spaces can be used to support different organizations of widgets.

Regarding visualization libraries, we chose the Google Chart library\(^9\), as it provides a convenient event system and it has a large support community. New visualizations are continuously being added.

In this iteration, we deploy seven widgets based on the visualizations from the previous iteration in iGoogle (see visualizations 1, 3, 5, 6, 7 and 8 in figure 1). Second, we changed the timeline (visualization 3 in Figure 1) by a motion chart (widget 2 in Figure 3). In the motion chart, x-axis is the activity of the user and y-axis is the peers average activity. A timeline chart can also be used to represent this data. However, when several goals overlap and are represented over the same time period, the user could be confused with too much lines and colors. A motion chart simplifies the representation.

Third, widget 3 in Figure 3 is added to the dashboard and centralizes filter information.

5.2 Evaluation

5.2.1 Evaluation data

Users were interviewed and observed during 15-30 minutes. They had to perform different predefined tasks such as filtering goals of this week following the think-aloud protocol.

In this iteration, we use hardcoded dummy data for the goal table and motion chart and data from a previous evaluation\([34]\). This allows us to emulate more realistic dashboard behavior.

5.2.2 Demographics and Evaluation setup

The digital prototype was evaluated with five male computer science teachers and assistants. Four of them are between 25 and 30 years old and one between 40 and 50. All of them know what a widget is. Three of them participated in the previous evaluation.

5.2.3 Evaluation results and discussion

One remark on this iteration is about our rationale (see Subsection 4.2.2) to create widget 3. The idea is to centralize all information in one place. However, the user perception is that every widget is independent from others, so they do not want to look up this information in another widget. In addition, depending on the screen resolution they have to scroll up to see the widget information and scroll down for the widget with the visualization. This dependency affects the usability of the application.

In this iteration, the selection of the table visualization receives good feedback due to the sorting functionality. However, regarding the calendar feature, users suggested to have different possibilities, such as a slider. In addition, functionalities to organize the goals by weeks and buttons for next and previous weeks were requested.

The motion chart was more complex than expected. Users spent quite some time using the different configuration options such as color and size. Although users consider it difficult to understand at the beginning, they indicated that the motion chart can provide useful information. However, all users remarked that they would like to see the dashboard in a real use case in order to assess its usefulness.

There are also minor remarks such as letter font and data inconsistency, small size of the text boxes and table filtering style.

We focus on solving the first issue. We propose to eliminate widget 3 and adding titles to the visualizations that can be updated based on filters. Removing this widgets also provides more space for bigger visualizations. For the next iteration, we eliminate calendar features because we do not need this kind of functionality in the use case study.

6. FIRST WORKING RELEASE

This iteration focuses on the real deployment of the dashboard. We selected an existing tracking system and adapted the existing widgets for this new scenario.

6.1 Design and implementation

We considered two tracking systems: RescueTime and Wakoopa\(^{10}\). They categorize tools and websites based on a functionality taxonomy such as Development, Browsers and Design. We selected RescueTime because it offers better security to access user data. As the next iteration with students involves real student data, such security and privacy considerations are very important.

We use the Rabbit Eclipse plug-in to track IDE Eclipse interaction. Students are developing software in Java and the Rabbit Eclipse plug-in allows tracking who is working

---

\(^{6}\)Open Social api, http://code.google.com/apis/opensocial/
\(^{7}\)iGoogle, http://www.google.com/ig
\(^{8}\)Apache Shindig, http://shindig.apache.org/
\(^{9}\)Google Chart library, http://code.google.com/apis/chart/
\(^{10}\)http://www.wakoopa.com
on which part of the project. The plug-in is open source and also tracks the time spent on documents (see figure 4).

The tracked information is collected via web services and exposed to the widgets via JSON. The dataset describes the time spent per application, document and website. This information is displayed in 8 different widgets as described in previous iterations.

In this iteration, we modify some widgets, because the RescueTime taxonomy enables us to categorize the tools by intended activity. This information can be useful for the students.

Widget 1 and 2 in Figure 5 are the same as described in Subection 5.1. Widget 3 shows the time per day spent by activity based on the taxonomy classification of RescueTime. The information is visualized using an annotated time line. Widget 4 is a bar chart that compares the global time spent per activity compared with the average time. Widget 5 shows the time spent per application. Widget 6 compares the time spent per application with other members of the group. Widget 7 shows the time spent on Eclipse projects files and, finally, widget 8 shows the time spent on websites compared with the average of the group.

The widgets use inter-widget communication for dataset filtering. Table 1 presents the connection details. This table explains which information is sent by every widget, and which widgets listen to events to filter their visualizations. For instance, when users click on a user in widget 1, this widget broadcasts the identifier of the user. Other widgets listen to this event and can show the information related to the user identifier.

**Table 1: Overview of events**

<table>
<thead>
<tr>
<th>Widget</th>
<th>Event</th>
<th>Listening widget</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>user identifier</td>
<td>2,3,4,5,6,7,8</td>
</tr>
<tr>
<td>2</td>
<td>goal identifier, start date and end date</td>
<td>3,4,5,6,7,8</td>
</tr>
<tr>
<td>3</td>
<td>selected range of time</td>
<td>3,5,6,7,8</td>
</tr>
<tr>
<td>4</td>
<td>type of activity</td>
<td>4 and 6</td>
</tr>
<tr>
<td>5</td>
<td>range of time</td>
<td>5 and 7</td>
</tr>
</tbody>
</table>
Figure 5: First release implementation
6.2 Evaluation

6.2.1 Evaluation data

In this iteration, we evaluated the dashboard with students. The data is tracked with RescueTime and the Rabbit Eclipse plug-in. As this evaluation took place at the start of the course, we did not have data from students, so two users (a developer and a project manager of our team) offered their RescueTime and Eclipse data for the experiment. The approach might influence the perceived usefulness, because students can not relate to their real data yet. However, the evaluation enabled us to obtain first feedback from students before the data collection started.

6.2.2 Demographics and Evaluation setup

This experiment ran with 36 students between 18 and 20 years old (30 males and 6 females) in an engineering bachelor course. We presented the dashboard the first day of the course. The privacy constraints and the data tracking characteristics of the experiment were explained. Students were also informed that they can stop RescueTime when they think it can affect their privacy. A questionnaire was used to collect quantitative data regarding first perceived usefulness, effectiveness, usability, satisfaction and privacy concerns. The questionnaire also has two open-ended questions about privacy considerations and general positive and negative aspects.

We wanted to evaluate whether students consider the dashboard useful and whether specific changes were needed to deploy the dashboard in this course. In the first question of the evaluation, the students get 80 points that they have to divide over the widgets to rank them. This question was intended to get insight into which visualizations are considered more valuable by the students. The next seven questions are extracted from the USE questionnaire [22]. The full questionnaire was not used due to time restrictions. Three questions are related to usefulness and effectiveness and the next four questions to usability and user satisfaction. Finally, the three last questions are related to privacy concerns. Question number 10 inquires whether students would be receptive to include tracking activity out of the lab.

6.2.3 Evaluation results and discussion

Results of the widget scoring question indicate that there is no clear winner (Figure 6), as we expected. Widget 3 and 4 have slightly higher scores. Both are related to activity type. Widgets 5 and 6 score the lowest. These widgets show the tools instead of activity type and can be found redundant.

Widget 8 provides information about what web sites have been visited and also scores high. In the open questions, 12 users find it useful to see what websites other students are visiting.

Widget 2, the motion chat, scores the lowest. Our perception is that the motion chart is more difficult to understand. In the next iterations, we pay special attention to the learnability of this visualization.

The questionnaire results are summarized in figure 7. The results indicate that students consider the dashboard useful (question 1 and 2) and that they think that the dashboard can help them to achieve goals (question 3). However, usability (question 4 and 5) and satisfaction (question 6 and 7) are scored neutral. As the students could not play yet with the dashboard, the scoring of these two factors was difficult.

From question 8 (see Figure 7), ‘I like to see what other members do during the course’, we learn that the students like to be aware of what their peers are doing. We can conclude from ‘I feel confident using the tracking system in the lab during the course’ (question 9), that the lab is a suitable context to track their activity. Question 10 (‘I would feel confident using the tracking system outside of the lab’) is rated the lowest. This outcome suggests that the students would feel uncomfortable if they were tracked outside the lab.

The open questions provide us with useful details. One of the most common remarks is that they like to see how students and their peers behave. 12 students mention that they like to compare their activity with others. However, 3 students indicate that they are disappointed that others can see their activity. One student suggests that tracking can cause stress and consequently decrease productivity [40]. Most students mention that the feeling of being observed is a negative aspect.
Another student argues that our visualizations can possibly modify their working style because they may want to behave similar to other students. We have to consider all these factors in future evaluations.

One student suggested to add support for detecting user distraction. Three users suggested to also enable access to the raw tracked data. These students were interested to see how RescueTime tracks data. Another proposal is to store the tracking information locally and ask for user permission before sending the information. This is an important suggestion to deal with potential privacy concerns.

7. SECOND WORKING RELEASE

The evaluation in the previous iteration is based on non-course data and a demo of the application. In this iteration, we focus on first results of the dashboard evaluation with real student data in a real course setting. As we describe in Section 8, more evaluations will be performed during the course in the following months.

7.1 Design and implementation

In this iteration, we analyzed the generated data to see how the students behave during the lab sessions. The dashboard was made available to 36 students. We created anonymous email and RescueTime accounts for each student. Students had to configure RescueTime and the Rabbit Eclipse plugin with their credentials.

7.2 Evaluation

7.2.1 Evaluation data

In this iteration, we evaluate the dashboard with student data. Students carried out different tasks, such as elaboration of scenarios, use cases and an implementation of a small web application during four lab sessions. As these tasks are partially performed without the computer, the tracked data is still limited in this phase.

7.2.2 Demographics and Evaluation setup

This experiment ran with 10 students between 18 and 20 years old (8 males and 1 females), a subgroup from the previous iteration. A subgroup was used to be able to better assist the students if problems would show up.

Students were encouraged to reflect on the dashboard visualizations during 10 minutes. They filled in a SUS questionnaire [10] afterwards. Such a questionnaire allows us to compare our application with more than 200 studies [7]. We added questions to score widgets (as used in the previous iteration), evaluate usefulness and satisfaction, and open ended questions.

During the evaluation, we removed widget 7 (see figure 5) because the only activity in Eclipse was the development during a tutorial, which would not provide useful information.

7.2.3 Evaluation results and discussion

The final SUS score is 72 points out of 100. Based on [7], this score rates the dashboard as good regarding usability.

The widget scoring question results are summarized in Figure 8. Widgets 4, 6 and 8 are rated highest. We think this is due to the limited data because we are in the initial period of the course. While bar charts display absolute information and are valuable even with limited data, timelines loose meaning because the user cannot see much evolution over the different sessions.

The 5-item likert scale, 'I would feel confident using it in another course', inquires about the usefulness and was rated on average 2.9. Users are not used to reflection on their own work using these tools. If the reflection task is mandatory during the course, they perform the task. However, they seem to prefer avoiding such tasks. The users do not find the dashboard beneficial enough to use it regularly. Part of this research is intended to increase the user’s interest for these kind of tools.

We asked the students about what they learnt. Three students highlighted the fact that there is not much data because they have not been working with the computers all the time. For instance, the dashboard does not represent the time students spent on the scenarios. Three other students found patterns in their Internet use. For instance, one student pointed out that his peers did not visit the course wiki as often as he did and realized that he was the person in charge to check this information. Five students indicated that visualizations are nice or even fun to use.

8. FUTURE WORK

The experiment runs during the whole semester and two more evaluations are scheduled. The essence of our future work is to actually evaluate the dashboard with the students as the course evolves and collects more real data. Such more elaborate data is required to assess in more detail the added value of these tools.

The current version of the dashboard enables students to compare their progress with peers on tasks that are defined by the teacher. In the next phase, we will add support to enable students to define their own goals. For instance, they could define how much time they want to spend every day on specific tasks. Self-definition of goals is an important part of Personal Informatics.

In addition, the dashboard technology allows easy customization. We can easily develop more visualization widgets that users can set up based on the context and their visualization background. We need to evaluate the influence of such customization factors in additional experiments.
9. CONCLUSION

In this paper, we presented the first results of a case study with second year engineering students.

We conclude that students consider the dashboard useful to learn how they are using the tools. However, users are not motivated to use the dashboard.

Visualization enables exploration of large datasets, but different visualization backgrounds can influence on the understanding of the data. Implementing the dashboard as a mash-up of widgets is our proposal to address this issue. The approach allows us to offer the users different visualization configurations.

The dashboard can be useful to support self-reflection and progress in comparison with peers. Students are interested to be aware of what their peers are doing. However, privacy concerns are involved in this process. The students are receptive to be tracked during their lab sessions. However, they do not like to be tracked outside a course environment due to privacy concerns. As additional work out of the lab sessions is not required for the course, this does not have implications on the current evaluation setup. However, the issue needs to be researched in order to generalize these kinds of experiments beyond the current course setting.

10. ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community Seventh Framework Programme (FP7/2007-2013) under grant agreement no 231396 (ROLE). Katrien Verbert is a Postdoctoral Fellow of the Research Foundation – Flanders (FWO).

11. REFERENCES


