

SESSION TRACKING FOR AN INTELLIGENT TUTORING SYSTEM

CARLOS HUMBERTO PAZ RODRIGUEZ
Msc. in Computer Science



Learning analytics

Software Engineering Department
Faculty of Science and Engineering

June 2018

ABSTRACT

Intelligent Tutoring Systems are gaining popularity for online learning. A recurrent problem is that teachers need more accurate monitoring tools that allow them to supervise and evaluate the work done by their students. Modern web tools provide functionality to detect user actions but they lack the capacity of correctly detecting continuous working sessions. A new activity tracking algorithm is proposed and implemented on a language learning platform named Zeeguu (<https://www.zeeguu.unibe.ch>). Accurate activity tracking increases reliability of data analysis and enables quality evaluation of the learning progress. The simplicity of the algorithm makes it applicable for areas even outside of the learning context.

ACKNOWLEDGMENTS

Many thanks to Oliver Holder, Christian Grier Mulvenna, Henry Salas, Tai-Ting Chen, Evi Xhelo, Ai Deng and Jakob Vokac for their contribution in the implementation of a teacher's dashboard that displays activity tracking data.

This work was supported in part by a grant from CONACYT / SICDET Michoacan, Mexico.

CONTENTS

1	INTRODUCTION	1
1.1	Background	1
1.2	Motivation	2
1.3	Outline	2
I	RELATED WORK	
2	LEARNING ANALYTICS	5
3	ACTIVITY TRACKING	6
4	DASHBOARDS	7
II	RESEARCH METHODOLOGY	
5	IMPLEMENTATION	9
6	INTERVIEWS	10
7	SESSION TRACKING	11
7.1	Algorithm	11
7.2	Reading session	12
7.3	Exercise session	12
8	DASHBOARD	14
III	RESULTS	
9	RESULTS	16
9.1	Reading session experiments	16
9.1.1	Baseline	16
9.1.2	Improving the baseline	17
9.2	Exercise session experiments	19
9.3	Dashboard	20
IV	DISCUSSION	
10	DISCUSSION	22
10.1	Reading session	22
10.2	Exercise session	23
10.3	Teacher dashboard	24
11	CONCLUSION	26
12	FUTURE WORK	27
V	APPENDIX	
A	INTERVIEW A	29
B	INTERVIEW B	33
	BIBLIOGRAPHY	36

LIST OF FIGURES

Figure 7.1	Finite state diagram - Working session transitions due to user actions	11
Figure 9.1	Reading session visualizations for session_timeout of 5 minutes.	17
Figure 9.2	Reading session visualizations for session_timeout of 2 minutes.	18
Figure 9.3	Reading session visualizations for session_timeout of 4 minutes.	19
Figure 10.1	Total time spent by a class, analyzed by date	22
Figure 10.2	Treemap of time spent by article. The color indicates the source of the article. Each block represents a user and the size represents the total time spent by a user on that article	23
Figure 10.3	Box plot of the average time required to read an article. It shows that most articles are read in less than 10 minutes and that there are some articles that take more time	23
Figure 10.4	Dashboard displaying individual statistics about reading habits	23
Figure 10.5	Class exercise time year over year comparison	24
Figure 10.6	Class exercise heatmap by weekday	24
Figure 10.7	Class average session duration	24
Figure 10.8	Class exercise time line	25
Figure 10.9	Bar chart showing exercise time by date	25

LIST OF TABLES

Table 9.1	User events classification, first iteration	16
Table 9.2	Average read sessions duration with different values of session_timeout	18
Table 9.3	Comparative of the average session duration using different session_timeout values for exercises	19

INTRODUCTION

1.1 BACKGROUND

Intelligent Tutoring Systems (ITS) are software solutions that monitor, evaluate and provide feedback to students about their learning behavior. They also manage the content to be studied and can provide some functionality to interact with other users. ITS are also used by teachers who monitor the students' progress and help them provide timely and personalized feedback during their learning process.

[8], [16] and [3] have used data mining algorithms to extract relevant patterns in the learning process. Typical approaches are implementation of association rules to discover relations between learning sequences or content. Clustering techniques are commonly used to sub classify a group of students into interesting categories and, for example, detect users with learning difficulties. Although some more advanced implementations thrive to automate the supervision and leverage teachers' decision process.

Ranging from normal statistical plots [16] to complex multidimensional visualizations [3], different approaches to provide understandable insight into the learning process have been researched. Histograms, pie charts and bar plots are well known visualizations that can be easily interpreted by untrained users, they are commonly used on dashboards and are useful for summarizing information. However, they are not powerful enough to display multiple evaluation criteria within a single screen.

Verbert et al.[20] has performed a comprehensive comparison of different analytical dashboards for ITS systems. He has classified them by target used (I.e.student or teacher) and by type of data being tracked (E.g.time spent, exercises results, documents interactions, etc).

Gamification [7] has also been proposed as a way to tackle lack of interest due to monotony. Motivation is a key factor in the success of a learning platform, especially due to the fact that there is not direct physical supervision. The student needs to be aware of the learning path and what distance he has traveled, thus realizing how much work is left to reach the desired goal.

User actions can be detected with Javascript events. Google Analytics is the most common tracking technology used nowadays for analyzing users' activity. However it only detects when a user enters and leaves a page, completely ignoring the effective time spent using the web page or if the user is browsing a completely unrelated website in a different tab of the browser [19] [12]. For the specific purpose of a web learning platform, it is critical to accurately track the time a user is investing on learning, thus knowing the time a user opened

and closed a text is most of the times not realistic. An algorithm to accurately detect effective working sessions is explained in chapter 7.

1.2 MOTIVATION

With the dynamic capabilities of the Web, a personalized textbook for learning foreign languages [10] is an ongoing research project. The tool allows users to read and learn new vocabulary by recommending texts of their own interest. A pilot test was implemented in a classroom. Both students and professors showed an interest in using 'the proposed tool. One of the outcomes from Lungu et al. research was that professors wanted to have monitoring functionalities that would allow them to keep track of their students' work and progress.

A major concern is that students can report false information regarding the work they have been doing with the application, and teachers' cannot provide timely corrective guidance. At the same time, students can have a misleading perception of the amount of work they have put on their studies, and might feel worried or unmotivated about their progress.

There are different theories over what is the best approach to learn a second language. Some theories say it is better to learn rules and grammar first, others state that practice is the most important element [5]. Kuppens has demonstrated that indirect exposure to a foreign language improves the knowledge compared to those who only study is during school. Therefore, the amount of time is an important metric to evaluate the progress done by a student.

Therefore the present research question is: How can I better measure the amount of time spent by a learner?

We know that with the continuous appearance of learning platforms and more accessibility to dynamic and free online content, it is certain that in the following years, new teaching methods that use technologies like this will emerge. As well, the results obtained with this research will be extensible not only to other learnings platforms but to any time tracking system.

1.3 OUTLINE

In Part I, state of the art work is discussed. First, an explanation of what learning analytics is and previous attempts to use activity tracking for this purpose are discussed. A specific gap is addressed that opens the need for this research project. Finalizing with an introductory explanation of what dashboards are and how they help in visual analytics.

In the second part, the research methodology is described. The goal of this project is to provide teachers who use Zeeguu with the tools to evaluate their students' work, therefore specific requirements were extracted by interviewing two key users. Once the critical requirements were obtained, a conceptual design of an algorithm for tracking users'

activity is described. Specific tweaks to the algorithm are used while implementing it for the reading and exercise activities.

After the code implementation, different tests were run to find the optimal settings for both implementations.

Finally, an analysis of the obtained results and possible conclusions are addressed.

Part I

RELATED WORK

LEARNING ANALYTICS

Learning Analytics is the research area focused on collecting, analyzing and visualizing information [2] related to the learning activities within a learning platform. Most of the current research focuses on Learning Management Systems (LMS) such as Moodle (<https://moodle.org/>) [13] [16] or Blackboard (<https://www.blackboard.com>) [1]. The existing research therefore focuses on systems that emulate a virtual university, which collects data about students, courses enrollments, assignments, exams and grades.

LEARNING MANAGEMENT SYSTEM Is the application that enables and manages all the educational content. Typically it also controls courses registration and tracks data about the students' progress. It also provides reports based on the collected data [21].

The nature of Personalized Textbooks implies that the learning road is different for each student, therefore the information collected for one user might not be the same for the rest (I. e. different subscriptions, the language being learned, the current level of knowledge, etc), making it more difficult to compare his learning performance against other students. Additionally, the level of detail is more specific than previous approaches. Other studies focused on analyzing grades and overall performance for different courses. The "personal" nature of Zeeguu, indicates that the information should be tracked for a specific article in a particular language. And because we want to track each individual reading or practice session.

In short, a new method for collecting and analyzing time spent on learning is required.

The term activity tracking has recently emerged and is mostly related to wearable devices that monitor fitness data [23]. In a more abstract sense, activity tracking is the process of collecting user's information over time for a particular purpose.

In the web context, technologies such as Javascript listeners and external tools such as Google analytics are combined to better detect when a user leaves a page. However it is impossible to track when a user closes the browser even for state of the art technologies, therefore Google analytics uses a configurable timeout parameter to close the activity session [14].

In a slightly different context, wearable devices [6] are not infallible under every external condition, and activity tracking can only be precise while collecting multiple measures and obtaining the average value [4].

Santos, Verbert, and Duval studied the effect of time tracking on the learning process. They concluded that time seems to be a good indicator of the learning progress of a student. During their experiments they used different tracking tools, from manual (E.g. Toggl) to automated (E.g. RescueTime) software applications. An advantage of manual tracking is that the student is more aware of his habits, therefore the commitment with the learning is strengthened, but the problem is that the students can declare false information, making it difficult for the teachers to detect problems and provide effective support. For automated time tracking Santos, Verbert, and Duval used RescueTime (<https://www.rescuetime.com/>), this tool tracks time spent on each application which differs with this project, where the interest is on accurately measuring reading and practice time. Manual tools require user input, Zeeguu's users are high school language teachers who want to prevent students from easily lying about their work. On the other hand, automated tools such as RescueTime, cannot really track text reading time, because a student can easily open an article, walk away and come back 1 hour later and pretend that they worked for 1 hour, while in reality they did not read at all. The most precise way of measuring reading activity is by implementing eye tracking solutions, which would result both expensive and invasive, and even though, it cannot detect if the user is really paying attention and learning or only staring at the text.

Other language learning systems such as Duolingo (<https://www.duolingo.com/>), only provide information about correct or incorrect exercises but no information about the time spent on the application, or time invested learning a foreign language. In chapter 5, a detailed description of a new approach for tracing reading sessions is provided.

DASHBOARDS

The term "dashboard" has evolved over time, the latest definition refers to a visualization of the most important information in order to achieve a particular goal, and within the visual space of a computer screen [13].

The main goal of a dashboard is to provide timely and useful information so that a decision can be made consequently. Verbert et al. describe the mental process that a user of a dashboard performs to make sense out of it. He states that a good dashboard should first help users acquire self-awareness by looking at data about their past activity and their current state in the learning process. Second, it invites him to do a personal reflection. So that the questions generated in the previous step get answered by exploring the visualizations and finally it must produce a change of behavior based on newly set goals. Telea describes the visualization design pipeline as an iterative process where the goal is to narrow the scope of the problem until a clear question can be answered with a visual tool. The visualization has to be designed for a specific user in mind, taking into account the complexity for interpreting the visualization and the level of precision when decoding the information back into abstract data.

A good quality dashboard should provide all the relevant information in a summarized way to enable the user solve a particular problem. Two set of principles were developed to evaluate dashboards design.

- SMART (synergetic, monitor KPIs, accurate, responsive, and timely)
- IMPACT (interactive, more data history, personalized, analytical, collaborative, and traceability) [11]

When creating a new dashboard, the most relevant information should be located at the center so that it stands out, and with the remaining space around the central visualization, it should provide details that support the answers obtained with the central element.

Part II

RESEARCH METHODOLOGY

IMPLEMENTATION

When implementing the session tracking algorithm, a correct set of rules need to be defined (E. g. what user actions are part of the active session, how much time a user takes to read a certain text, how much time can a user read without translating, etc) in order to properly measure a working session.

Two language teachers, one teaching French and the second teaching English, have been using the Zeeguu platform during their lessons, therefore they were the ideal candidates for discussing about how to properly measure their students' work. In chapter 6, the structure of the interview is explained.

Once obtained the general outlines, the algorithm design was started. The algorithm was designed following the finite state machine approach, where the session lifespan transitions between different status until the closure is reached. The technical details of the algorithm are addressed in chapter 7.

Finally, a dashboard was developed to enable teachers evaluate the results of the session tracking (chapter 8).

6

INTERVIEWS

During the first pilot test of the Zeeguu system, users demonstrated an interest in continue using the system. However, the two teachers that used the platform concluded that is difficult to keep track of the time spent by the students. One of them decided to ask his class a weekly report mentioning the amount of time and the read articles. The second one, decided to evaluate only the words that they translate, since this was the only available information.

Same goes for exercises, the results can be given but not how much time was spent by the students reinforcing their knowledge.

During the interviews, both teachers mentioned the need of precisely measuring the reading and practice time. The first teacher grants his students points for every half hour of work, which puts an extra pressure on the students, and sometimes they reported more reading time than real. This happens specially when students are reaching a deadline and have not made enough work.

One teacher also evaluates the amount of work every week and gives them a classification (I. e. green, orange and red), which enforces the commitment of students.

A transcript of the interviews can be found in the Appendix section (v).

SESSION TRACKING

In order to understand the learning process, we need to keep track of the working sessions. The term refers to either reading or exercising sessions. A reading session is considered as an uninterrupted reading activity of a particular article. A similar concept applies for exercises, except that they can encompass multiple articles. Given the fact that students can switch between windows or even walk away from the screen, a method to compute the effective working time is devised.

7.1 ALGORITHM

The algorithm to compute the reading session works under the assumption that all user actions fall under one out of three types: opening, interaction or closing events.

By following the Markov model on figure 7.1, we observe that depending on the type of event, a working session can be either created, updated (kept alive) or closed. With this simple approach, the algorithm can be fine tuned by adding or removing events (E. g.including the scrolling as a interaction event).

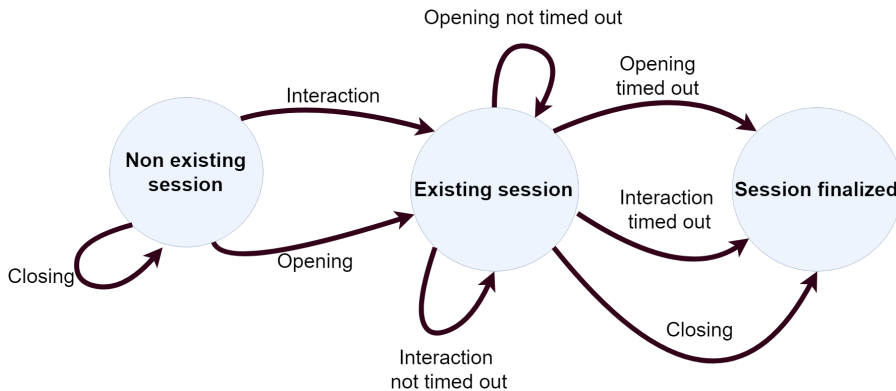


Figure 7.1: Finite state diagram - Working session transitions due to user actions

Additionally, there is a **session_timeout** parameter which is used to handle scenarios where a session was not properly closed (E. g.when the browser was closed or the computer was turned off). The parameter is used in 2 different situations:

- When the user is inactive for a long period, the session is closed and a bonus of "+ session_timeout" minutes are added to the last action date and time. Under the assumption that the user did not stop working immediately after the last action.

- When the user is taking unusually too long between actions the activity is considered as suspicious and the session is closed. A new session is created with the next user action.

7.2 READING SESSION

With the help of Javascript code, user actions like opening an article, translating a word, or even switching the browser tab can be detected and tracked, however events like when the user abruptly closes the browser or turns off the computer, or even when the user walks away from the screen cannot be detected using Javascript. For these scenarios, a timeout parameter is used that finalizes the session after **N** number of minutes. If the timeout value is too small an advanced student that does not translate any word but reads a long text might be incorrectly measured, while if the value is too big, we would incorrectly consider longer reading sessions.

Some important assumptions for the implementation of the reading session are:

- Reading sessions are considered per article. Therefore, when the user closes and opens a new article we consider it as a new session. In this way we solve an issue where the same user could have open in different tabs different articles and switches between them.
- For scenarios when the user leaves the session open for a time period longer than the `session_timeout`, we give it a time benefit of “`session_timeout`” minutes, because we cannot know exactly how long after the last action, the user kept reading.
- When the user opens a new article without closing the previous one(s) we close the other reading session(s) because the user is supposed to read one article at the time.
- A user can read the same article multiple times, and each time is considered a new session (unless the time between both of them is less than the session timeout, in which case the two sessions are merged).

7.3 EXERCISE SESSION

In the case of exercise sessions, the implementation is simpler, because the only possible interaction is answering either correctly or wrongly the exercise.

Therefore, opening an exercise does not count for the active session, only answering. As long as the time between two exercises is smaller than the `session_timeout` the exercises will be placed under the same exercise session.

For exercise sessions we can be more precise on the tracking because an exercise usually comes surrounded by a single sentence, therefore the timeout can be set to a smaller value.

The final difference with the reading session is that no "benefit time" is granted to the user, the only actions that keep the session alive are answering the exercise.

As mentioned earlier, learning analytics is about providing information to the user so that they can reflect about their own activity and modify their actions to better achieve the expected results. For this reason a visualization of the session tracking is implemented.

By analyzing the feedback provided by the teachers during the interviews, they expressed that they would like to view the activity of their class but mostly at the individual level, given the personal nature of the Zeeguu platform.

A simple dashboard with a heat map was designed, providing a visual overview of the class activity in a single screen. The teacher is therefore able to view which students are active and when, what type of activity they do and for how long. A heat map was chosen because is an easy to interpret visualization that does not require any particular expertise (figure ??).

Additionally, a detailed view of how individual work is decomposed is provided by means of simple bars and percentages that give the teacher information about how much time was spent on reading and how much on exercises (figure ??).

Interactive functionalities are also included, so that the teacher can analyze the class on his own and make particular discoveries (E. g.what students are falling behind, who has not worked in the last week, etc).

Part III

RESULTS

RESULTS

In this section, a series of iterative experiments were performed to find the best implementation/settings of the algorithm. As explained in chapter 7, a wrong value of the "session_timeout" parameter can produce different and perhaps incorrect results.

9.1 READING SESSION EXPERIMENTS

The reading session tracking is the most complex to model, because there are multiple internal and external variables to consider. First, we must define what actions are relevant to detect, in order to consider them as part of the learning process, and which ones are not (I.e. the user trying to game the system). Second, with the events already defined, we must determine how much time can pass between them, which is still part of the same reading/learning session.

9.1.1 Baseline

For the first iteration, the initial set of **opening, interaction and closing events** were defined as listed in table 9.1.

EVENT	ACTION TYPE
UMR - ARTICLE CLOSED	Closing
UMR - CHANGE ORIENTATION	Interaction
UMR - CLOSE ALTERMENU	Interaction
UMR - OPEN ALTERMENU	Interaction
UMR - SEND SUGGESTION	Interaction
UMR - SPEAK TEXT	Interaction
UMR - TRANSLATE TEXT	Interaction
UMR - UNDO TEXT TRANSLATION	Interaction
UMR - ARTICLE FOCUSED	Opening
UMR - OPEN ARTICLE	Opening
UMR - OPEN STARRED ARTICLE	Opening
UMR - ARTICLES REQUESTED FROM ZEEGUU	Closing

Table 9.1: User events classification, first iteration

For the **session_timeout**, the initial test was done with 5 minutes.

By observing the total reading time by user, we can see that the highest value is 2120 minutes (roughly 1.4 days) and an average time of 11 minutes spent by the user with the longest average reading time. Additionally, by analyzing the daily activity for a specific user,

we find long reading sessions spanning up to 1:20 hours long. Figure 9.1 show these results.

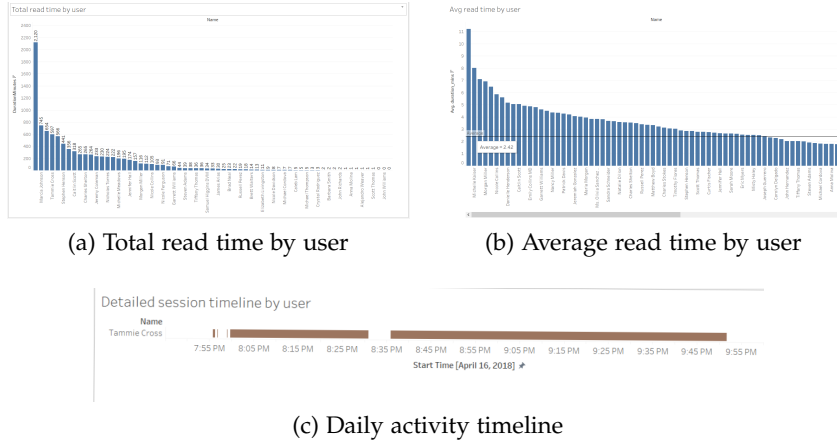


Figure 9.1: Reading session visualizations for session_timeout of 5 minutes.

9.1.2 Improving the baseline

Using the previous settings as a baseline, the next idea was to reduce the value of session_timeout, because the additional time benefit of 5 minutes might be too much of a consideration for a user reading without doing any action. For that to be possible, **scrolling events** were included as interaction events, so that they keep the session alive. Scrolling events are helpful for detecting more user activity, but they can be overwhelming for the web server and the database, for that reason, the event posting was limited to maximum 1 per second.

One of the interesting findings in the first iteration is the mean of the average read time for all the users (figure 9.1b), we can observe that the average reader works for about 2.42 minutes. Taking this discovery into consideration plus scroll detection, we reduced the value of session_timeout to 2 minutes and ran a second iteration of experiments.

By observing figure 9.2a and comparing it to the results obtained in figure 9.1a, we find a clear change in the total reading time by user, the largest amount was 2120 minutes while now the largest is 1009. A similar behavior is observed for the rest of the observations, this tells us that the previous session_timeout value was too big, therefore computing times longer than what they really were.

For the average reading time by user plot, we observe that the values do not change much from the previous results (figures 9.2b vs 9.1b). The longest average time is now 9 minutes instead of 11. However, now the mean of all the values moved up to 2.6 minutes, which tells us that the value of 2 minutes is smaller than what we need.

Finally, for the detailed activity, a big change is observed, in figure 9.1c, we observe roughly two smooth reading sessions, which now with a finer session detection broke down into smaller reading sessions (figure 9.2c). This means that we correctly detected the

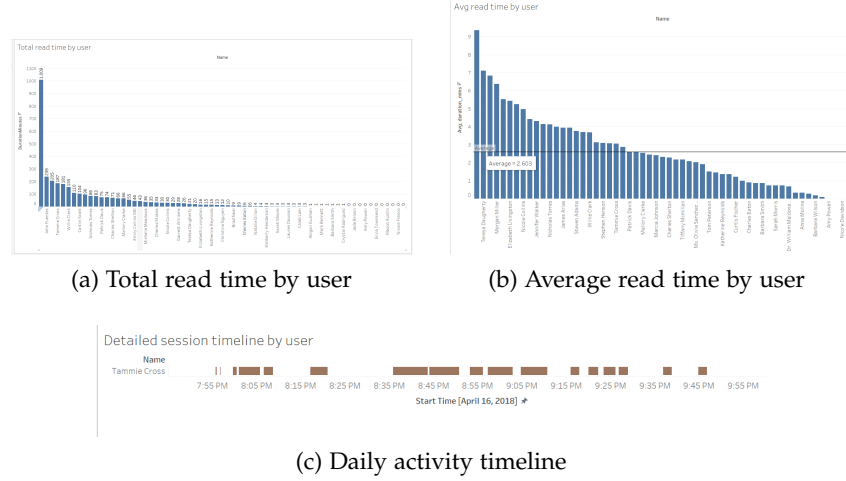


Figure 9.2: Reading session visualizations for session_timeout of 2 minutes.

user's gap of attentions, therefore a more precise session time was measured.

After the initial findings, the session tracking calculation was run with different parameters to find the best settings. In table 9.2, the average session durations are compared.

SESSION_TIMEOUT VALUE	AVG. READ SESSION (IN MINUTES)
1	1
2	1
2.6	3
3.2	4
4	4
5	5

Table 9.2: Average read sessions duration with different values of session_timeout

As we can observe for the average session time, for low values, the average session duration decreases because the sessions are being split into smaller chunks. And for high values, the small sessions are stitched, thus moving the average to a higher value. However, for values between 2 and 4 minutes, we observe that the average stabilizes, therefore we use the value of 4 minutes as the final reading session_timeout.

By running the reading session computations with the session timeout of 4 minutes, we observe an increase in the total time by around 30% but still not as much as the baseline (figure 9.3a). The average reading time stabilizes at 4 minutes (figure 9.3b). And the detailed time line sessions end up in an intermediate outcome between the two previous examples (figure 9.3c).

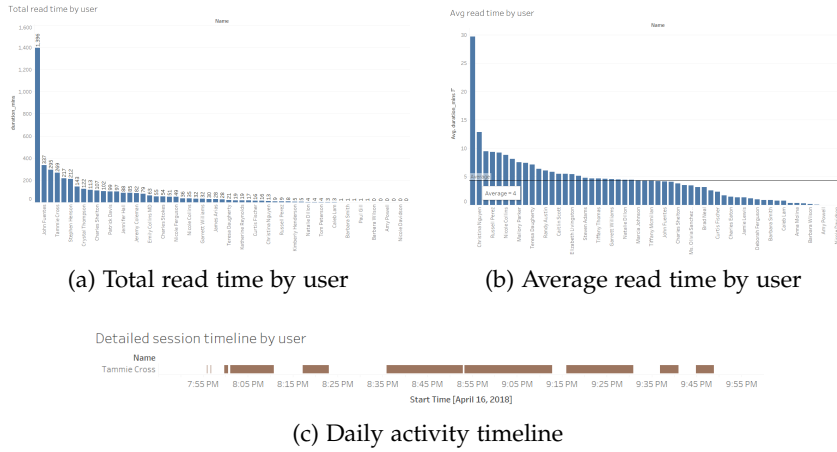


Figure 9.3: Reading session visualizations for session_timeout of 4 minutes.

9.2 EXERCISE SESSION EXPERIMENTS

For the exercise sessions, the implementation is quite simple, the only event that we track is answering of the exercise (either right or wrong or event asking for a hint). Therefore the only parameter to define is the exercise session_timeout. For the exercises, we know that they are short and quick activities, which take only a couple of second to provide an answer. Therefore the timeout must be much smaller than for reading.

Again, a series of experiments were run with different settings to detect the optimal timeout value. In table 9.3 we can observe the average exercise session length computed for different parameters of the session_timeout.

SESSION_TIMEOUT VALUE	AVG. SESSION (IN MINUTES)
1	2.7
2	3.7
2.5	4
3	4.2
4	4.7
5	5.22

Table 9.3: Comparative of the average session duration using different session_timeout values for exercises

Based on results, it is difficult to find a clear winning value, however, fluctuations decrease around a value of 4 minutes. Of course, if we keep increasing the value, we would find longer sessions, but since we want to be as strict as possible, a value of 4 is chosen as the optimal.

9.3 DASHBOARD

TO BE DEFINED

Part IV

DISCUSSION

After the implementation of session tracking, useful information about time and quality of work can be presented to teachers and students. Example of questions and analysis that can now be performed with the session tracking are presented in this section.

10.1 READING SESSION

After measuring the time spent by the students, useful information can be provided to the teachers, with this knowledge they can adjust their teaching strategy and have a better understanding of how the class is using the system.

By visualizing the amount of time spent for the students in a class, we can detect "islands" of activity common for the whole class (figure 10.1a). These so called "islands" can be explained with a visualization by week day. In figure 10.1b we can confirm that the students are mostly actively reading on Sundays and Mondays.

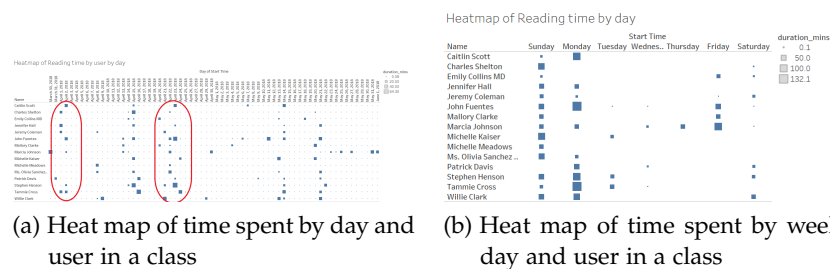


Figure 10.1: Total time spent by a class, analyzed by date

Since the reading sessions are associated to an article, an analysis of the time spent by article can be performed. In figure 10.2 we display the total amount of time spent by a class of students learning Dutch. With this information the teacher can observe which articles are the ones that most students have read, and also from which source they read, so a further discussion in class about the content can be done.

An alternative way of grasping an idea about text difficulty and quality of work can be done based on the time students spend on average reading an article. In figure 10.3, we can observe that most of the articles were read during less than ten minutes, amongst which the majority are concentrated in less than 4 minutes sessions.

And at the individual level, the user can visualize and understand his own work, in order to be aware of his own progress. In figure 10.4, an example of a dashboard containing the information that could be relevant for a user is displayed.

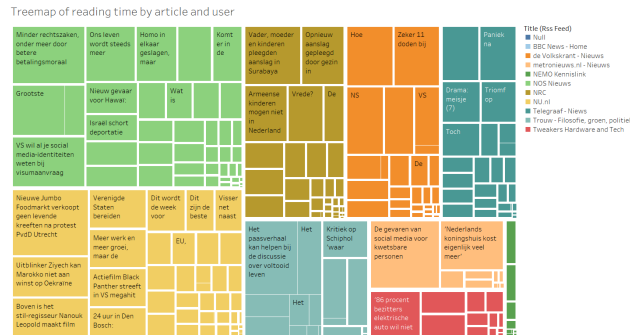


Figure 10.2: Treemap of time spent by article. The color indicates the source of the article. Each block represents a user and the size represents the total time spent by a user on that article

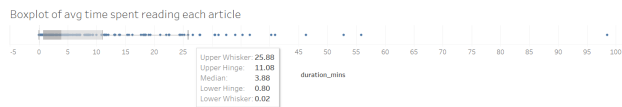


Figure 10.3: Box plot of the average time required to read an article. It shows that most articles are read in less than 10 minutes and that there are some articles that take more time

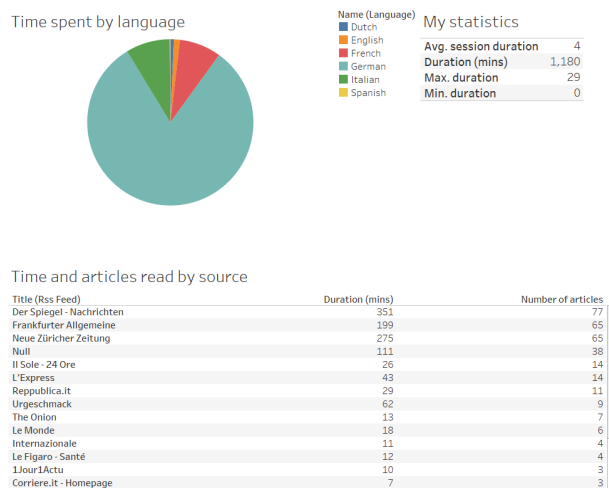


Figure 10.4: Dashboard displaying individual statistics about reading habits

10.2 EXERCISE SESSION

Typical questions related to the exercises are how much time has the student being practicing? Do they practice daily? For how long? Etc. In this section a list of analysis based on time spent on exercises is shown and discussed.

The first question is how much time the class is spending on practicing their vocabulary? With a simple bar chart we can visualize and compare how much time in total has the class spent practicing, and also who has been putting more effort on it. As we can observe in figure 10.5, some students really like to rehearse their new vocabulary, while the rest have lost their interest over time.

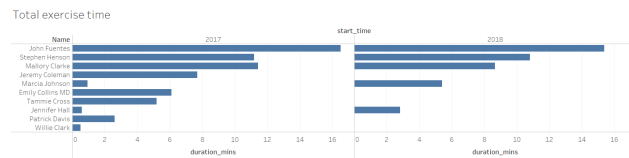


Figure 10.5: Class exercise time year over year comparison

If the interest is on when do the students practice, a visualization by weekday can show us that, for example, for a particular class, most of the students prefer to practice on Wednesday (figure 10.6).

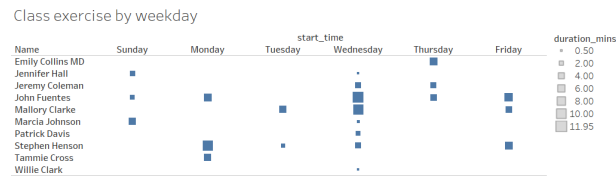


Figure 10.6: Class exercise heatmap by weekday

For understanding the quality of the practice, we can also visualize the average session length. An interesting finding, as we can see in figure 10.7, is that on average the exercise sessions are 2.2 minutes long. Some of the students however practice for less than a minute, this can tell us about their commitment with the activity. By observing these results a teacher can encourage the class to practice for at least a certain amount of minutes in a row, in order to keep a better focus. Finally, this visualization can also be used to avoid students gaming the system by opening the application multiple times but not really focusing on the assignment. The finding about the average session time is also a good indicator about how the exercise session_timeout should be configured.

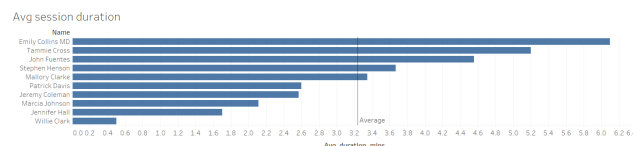


Figure 10.7: Class average session duration

Continuing with the analysis of the quality of the work, a detailed time line (figure 10.8) is a more fine grained visualization, where a teacher can observe at what time of the day the students practiced and if they worked uninterruptedly or not.

As well, from the individual perspective, a student can observe how is his progress and consistency over time. He can set himself personal goals and track over time how is his performance. In figure 10.9 we observe that the student was not constant in his daily practice.

10.3 TEACHER DASHBOARD

TO BE DEFINED

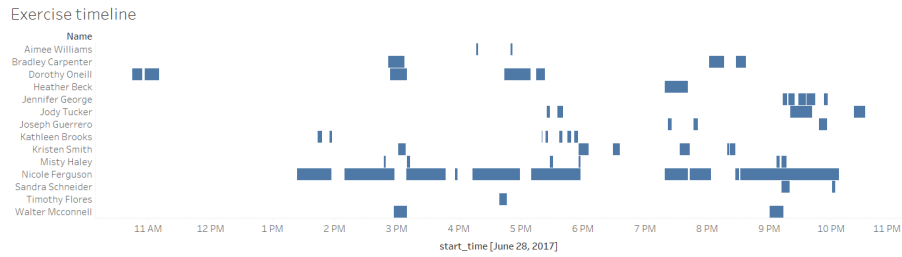


Figure 10.8: Class exercise time line



Figure 10.9: Bar chart showing exercise time by date

As observed in section 9, not only the time spent but also the quality of the work can be assessed with a precise session tracking. A teacher can use multiple approaches to evaluate the progress of students in the individual level and also as a class.

The extensibility and applicability of the session tracking algorithm is vast. The proposed algorithm can be used for other purpose ITS, by only adjusting the timeout and the particular set of rules for the specific application. As well as for the commercial industry, the session tracking algorithm can be used to track customers engagement with a particular website, and interesting questions can be crafted and answered via visualizations. In the office, session tracking can also be implemented to detect employees who are having too much distraction from their activities.

For our particular application, the Zeeguu platform, one of the main goals was to keep students from gaming the system. With that goal in mind the event tracking and the timeout parameters had to be strict. However, not for every application must be the same.

Finding the optimal timeout value is not a trivial task, it requires sample data and multiple tests to find a good setting, and even then, it can vary from context to context (E.g. depending on the level of expertise of users, on the type of ITS, on the level of trust that can be placed on users, etc).

FUTURE WORK

One of the main challenges for defining the `session_timeout` is having statistically representative data. In this case, Zeeguu system has been recently deployed and has suffered constant modifications (being one of the latests scroll and page focus detection), this skews the analysis to only the last collected information. For this reason it is possible that in the future, with more data at hand, the `session_timeout` can be better adjusted.

A possible improvement to the implementation is having different `session_timeout` values, either by class or by student. Helping better tailor the parameters and obtaining even better results. In the ideal system, if we knew that an article was too difficult for a user, we would expect a lot of translations, therefore the timeout could be set to a smaller value. On the contrary, for advanced users, the timeout could be larger given that they do not need to translate any word.

Once enough data is collected, reading speed calculation can be done, which could be used for filtering suggested articles based on available time, or even to make the session timeout smarter.

The next step for the Zeeguu system will be moving to data mining and predictive analytics. For example, based on historical activity, the system could evaluate and categorize students based on their expected performance, and perhaps raise an alert for those who are under performing.

Part V

APPENDIX

INTERVIEW A

- *Question: What kind of information would you like to get and what is the purpose?*

I give a lot of freedom to my students to choose their activities. I don't mind what they do as long as they are reading. That is a very modern approach to language teaching so they are exposed a lot to the language, and the only thing I want is that they spend for example half an hour a day or an hour a week on reading, and I award them with points for every half hour of work. They have to get 60 points a year for example. They get to choose for reading activities, speaking activities or writing activities. When it comes to reading activities that is when Zeeguu comes in. A number of students are very fond of Zeeguu, they don't do anything else for reading, so they spend a couple of hours or minutes on the system reading, checking for translations, doing exercises. My students are very motivated but even motivated students, when the deadline approaches tend to become less motivated to do the things right away. They tend to become a bit more easy on the honesty, so when the deadline approaches they might be doing something during 10 minutes and they report on the paper that they spent 60 minutes. I say to my students I cannot look into your head, I cannot see what's happening in your head, I cannot see what you're doing in terms of language acquisition the only thing I can see is that you spent an amount of time on reading activities. That is where this dashboard might be a little more helpful than now, because now I only have their words, so they fill out the report saying for example that they spend 30 minutes or 60 minutes on a certain text I can look into the system and I can see the difficulty of the text and I can see how many words they have asked for a translation but that's all. So what I need is more things that give me the idea that they really worked for example 30 minutes. I need to be convinced that they are doing that amount of time. I don't see anything on their activities like their exercises, the time if that is possible. I don't think it's fair when I don't convince my students that I have serious instruments that can make a positive guess about their work. When I don't do that I leave it to their self-discipline and I don't think it's fair to the students. They need to have some kind of guidance, so I have to put a little pressure on them to do the right thing.

- *Question: So that reinforces the idea that both you and your students can see the same information and they can realize that this is what the professor is going to visualize, so maybe I have to put a little extra effort?*

Yeah that would be very nice.

- *Question: Talking about the weekly report that you mentioned, you only ask them how many hours they worked?*

They fill out of this form telling me their name, the classroom, the date that they did this activity, the text title, the magazine they chose the text from and how many minutes, 30 or 60 minutes.

- *Question: I'm just trying to understand the purpose behind the report, one side is for keep track of what they are doing but is it also for motivating?*

Yes it's external motivation, because most of students don't like to do homework when they are home they want to do other things. I have to put pressure on them to do the things they really should do in order to get to the exam. so I have chosen for a system in which I can rely on them to do things at home by giving them 60 points for 30 hours, that means about one hour a week, and as a personalized system they can choose whatever activity they want to do. In the end, when they don't have the 60 points they can't go to the next year.

- *Question: The next question is how will you measure the quality of their work? Would it be only the amount of time do they spend reading? Or as well the score on the exercises? Or do you have another way of evaluating the quality?*

Now I only have quantitative measures I can look at the text, I can look at what words they have been asking to be translated and that gives me an idea about the quality of their work. Sometimes I see that they ask for too easy words and I talk to them about it. Whenever I see something that I don't like I write down the names and after class I ask them to come, and I ask them "why did you look up that word? You know this word, or at least I think you should know this word". From time to time, I have a discussion with those students to probe quality, but that is difficult because when I say "I think you have learnt these words in the first class 4 years ago", they can say "well, I forgot" and there is nothing I can do about it. I keep asking those questions and when I ask those questions that is a way I put pressure, they know that if they take it too easy I can ask a lot of questions that they don't like so they are triggered to do their best I think. But it might be helpful indeed to have some information about how they performed in those exercises.

- *Question: And after how many correct exercises of a specific word would you say that they have learned the word? We can ask them ten times the same word and maybe after the fifth time they have learned the word already.*

That is a very difficult question; the basis of the approach that I use is statistical learning. They see words, they see texts, they

see language, and they make inductions about that, and every time whenever they see it for the second, third, fourth time it is enhanced, that meaning is best kept in mind, so every repetition is welcome. It's a bit compared to when you learn your mother tongue you have a lot of exposure to the language, the amount of exposure is so enormous to the first language and I don't get that the amount of exposure with my students because I only have three classes a week.

- *Question: Are you interested in analyzing individual students, or as a class or both?*

Individual in the first place because it's personalized, but to have information about how a class performs might be helpful as well. I have for example two classes one of them is very loyal, always have the points on time and the other group doesn't. That is something that I see and this will be reinforced by information; but for me it's not necessary because I evaluate them on the individual level. And the second problem is that not every student uses Zeeguu to get his points.

- *Question: Another question will be what level of detail are you interested? For example at the word level or a bit more general for example at the session level, and how many sessions or how much time have they spent?*

The amount of time they spend on the system is very important, I think. And the amount of hours they do during the exercises might be a good indicator of their effort. When they are having a lot of errors, they are not having a mental activity they are just choosing whatever. When they are busy reading they are busy with meaning.

- *Question: How often did you evaluate their work? Is it a daily, weekly, monthly?*

Two monthly. I have five times a year so I have 5 deadlines and that is in terms of 2 months. I asked for example 12 points by the end of the first two months. And if they don't, they get an application with me to do two points a week. I work with levels: orange, green and red so when you are on level green you're ok, you can choose whatever you want, and when you don't meet my demands you go to level orange. That means every week you have to produce two points. And after the second time that they don't reach the deadline, they go to level red.

- *Question: This question is related to analyzing the students as a class. Are you interested in detecting special students like the ones that are working a lot?*

It doesn't tell me much because some students don't use the system. What might be interesting is that the students might be able to see the ranking of the articles. The ones that they like a lot. Maybe that can be done in the system. It saves a lot of time

to them because I have to look at the text and decide which one to take. And with a lot of extra information from other students, it might save time.

- *Question: Is there something that you would like to add?*

Normally teachers give homework that is not personalized. Students have text in the book for example or text from the internet that is copied and they work with that. What might be a good addition for these colleges for example is that the teacher chooses a text and designs a sort of test to be able to verify that they have done the work and that is of quality. I am not really interested in quality but in quantity.

INTERVIEW B

- *Question: As a teacher, what would you like to visualize or to understand from the system?*

I would like to know more or less what kind of resources they are using and what are the kind of articles they are reading. Let's say at this moment that a student reads a text about the sun and another student read that text about the same topic in the Guardian. It will be great to have a visualization on the resources, I want my students to read The Guardian for instance, but I cannot force them to do so, but it will be great to have a clearer view of the resources they use, and if possible the kind of article. imagine that these articles are already categorized according to the newspaper, so one article is about politics, and others about sports or something, it would be great to have something as well in the tool. So that you can easily say, I want you to read a text about sports but you didn't because you read something about politics. So some categorization of text types and resources. The time students spend actually reading, that is I think already automatic in Firefox you have the reading tool in Firefox and it automatically says how much time you should spend approximately on reading that article. Let's say you are reading an article in the Guardian and it says, well it's approximately 6 minutes. But then it would be great to have it measured if it's really six or seven minutes. It will be great to see how many words a student looks for. Imagine that you have a text of 200 words and that student searches for 150 words, then that information could help you help the student. The amount of time students spend on the exercises, because now it says two minutes three minutes or five minutes which will take approximately. You can more or less say you have one hour for that week and that means that I have 12 texts of 5 minutes and I want you to read this texts. And the difficulty of the text, so if a student reads only the more simple texts then you'll see it's not improving, you should more or less see some kind of improvement over time. Now the tool is very user centered, so is a student who reads text autonomously, and as a teacher, you could say ok I want you to read so many minutes, so that the system collects this information. Eventually it might also be interesting to be able to limit the amount of texts to say I want you to do this or that, as a teacher.

- *Question: How would you measure the quality of the work?*

It depends on the sources, on the text length (the readability score). I see the activities afterwards like exercises more as a way to practice, a fun way to work with words. But I find a text more

important, the text complexity and the way they interact with text. I see the activities afterwards as an attractive way to spend a little more time on the words that I looked up but it's not the best way, you can also use a dictionary or write the words in their own dictionary. I personally feel that they have to do more with the words. You can also say I just want to improve my receptive word knowledge which is also great, but on the other hand, you also want to use these words, that is quite complex to interact with a text that you would normally not interact with. That would also be interesting to see how to force people to interact with the text. I could be doing a summary or giving an opinion.

- *Question: Is it important for you to understand each student individually or as a class?*

Preferably, each student individually, but it would also be great to have an overview of the class as a group. The common European framework for reference it's a bible for people who are familiar with the language portfolios, Zeeguu can be seen as a repository for demonstrating your reading knowledge. You are collecting online text and that tells about your reading knowledge. Therefore you can then say that a text is for instance at an A2 level, but imagine a text about environmental problems in America will be more at a B2 or C1 level. So that means that I am improving myself, so eventually I am in this level.

- *Question: How would you categorize a text according to a different level in the framework?*

According to the framework you can only read texts from B1 and so on. But then again that is not entirely true because you can use translation tools in Zeeguu, so Zeeguu allows you to read texts even if you are not B1 or B2. Imagine that you have a frequency, and 80% of the words of a text is in the range from 0 to 2000, so if you now these words, which you should already know at a A2 level, you should be able to search for the difficult words in the higher range meaning that you can basically read an authentic text because you already know the other kinds of words, and there is information about the frequency list.

- *Question: How would you classify a word as known or unknown? If you find the text and you don't translate any word do we infer that they know all the words?*

That is quite difficult so maybe you should try to avoid that question, that is something that we discussed about, the question is that if they improve overtime? That is really difficult to answer. For a learner if you use the tool and it works for you, that's great. And as a teacher you can ask your students to use the tool to improve their language, but maybe that's all.

- *Question: When you evaluate your students, what kind of time window do you use?*

Probably weekly.

- *Question: Is it important for you to detect special cases of students? Like outliers from the normal class? Or do you prefer to analyze them individually?*

It is quite difficult what you propose, imagine a student that reads only texts from the sun and does not look up for any word, and you have students who read text from the Guardian who look up for 20 words because they have intrinsic motivation to improve their words, so it is difficult to compare different articles.

- *Question: With your current experience with the system, how do you evaluate your students? Or their progress?*

I see my students every week so I talk to them about the tool, and I made it obligatory for them to use the tool in order to improve their reading knowledge. I have my reader with seven thousands words, they are forced to study these words study these words, I have the same approach. I have to first look into the type of word (is it a verb, or a noun, or an adjective and so on), they probably have to look up for some words and then they have to come up with ten questions that they have to ask the other students, and they have these words in brackets as a stimulus for response. So I want them to use these words in different circumstances. That is explicit, then Zeeguu as a fun way to learn words implicitly, so I have two different approaches to vocabulary.

- *Question: Did I miss something that you consider relevant?*

It would be great to have more information about time that students spend, about the reading time, text complexity, resources and perhaps also text topic. I was just wondering if it would be possible to select texts based on difficulty as well, according to the measurement tool that you are using right now. Imagine that one wants to read only texts with a difficulty of 8 or above in the selected languages, because less than 8 is supposedly too easy.

BIBLIOGRAPHY

- [1] Kimberly E. Arnold, Matthew D. Pistilli, and Kimberly E. Arnold. "Course Signals at Purdue: Using Learning Analytics to Increase Student Success." In: *2nd International Conference on Learning Analytics and Knowledge* May (2012), pp. 2–5. ISSN: ISSN-1528-5324. DOI: [10.1145/2330601.2330666](https://doi.org/10.1145/2330601.2330666). arXiv: [0709.1706v2](https://arxiv.org/abs/0709.1706v2).
- [2] Malcolm Brown. "Learning Analytics: The Coming Third Wave." In: *EDUCAUSE Learning Initiative Brief* April (2011), pp. 1–4. ISSN: 2150-6000. DOI: [10.1007/978-1-4614-3305-7](https://doi.org/10.1007/978-1-4614-3305-7). URL: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Learning+Analytics:+The+coming+third+wave{\\#}0>.
- [3] Buket Dogan and A. Yilmaz Camurcu. "Visual clustering of multidimensional educational data from an intelligent tutoring system." In: *Computer Applications in Engineering Education* 18.2 (2009), pp. 375–382. ISSN: 10613773. DOI: [10.1002/cae.20272](https://doi.org/10.1002/cae.20272). URL: <http://doi.wiley.com/10.1002/cae.20272>.
- [4] Fatema El-Amrawy and Mohamed Ismail Nounou. "Are currently available wearable devices for activity tracking and heart rate monitoring accurate, precise, and medically beneficial?" In: *Healthcare Informatics Research* 21.4 (2015), pp. 315–320. ISSN: 2093369X. DOI: [10.4258/hir.2015.21.4.315](https://doi.org/10.4258/hir.2015.21.4.315).
- [5] Nick Ellis. "Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge." In: *European Journal of Cognitive Psychology* 5.3 (1993), pp. 289–318. ISSN: 0954-1446. DOI: [10.1080/09541449308520120](https://doi.org/10.1080/09541449308520120). URL: <http://www.tandfonline.com/doi/abs/10.1080/09541449308520120>.
- [6] Kelly R. Evenson, Michelle M. Goto, and Robert D. Furberg. *Systematic review of the validity and reliability of consumer-wearable activity trackers*. 2015. DOI: [10.1186/s12966-015-0314-1](https://doi.org/10.1186/s12966-015-0314-1).
- [7] Carina González, Alberto Mora, and Pedro Toledo. "Gamification in intelligent tutoring systems." In: *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality - TEEM '14*. New York, New York, USA: ACM Press, 2014, pp. 221–225. ISBN: 9781450328968. DOI: [10.1145/2669711.2669903](https://doi.org/10.1145/2669711.2669903). URL: <http://dl.acm.org/citation.cfm?doid=2669711.2669903>.
- [8] I. Jugo, B. Kovačić, and V. Slavuj. "Using data mining for learning path recommendation and visualization in an intelligent tutoring system." In: *2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2014 - Proceedings*. 2014. ISBN: 9789532330816. DOI: [10.1109/MIPRO.2014.6859700](https://doi.org/10.1109/MIPRO.2014.6859700).

- [9] An H. Kuppens. "Incidental foreign language acquisition from media exposure." In: *Learning, Media and Technology* 35.1 (2010), pp. 65–85. ISSN: 17439884. DOI: [10.1080/17439880903561876](https://doi.org/10.1080/17439880903561876).
- [10] Mircea F Lungu, Luc Van Der Brand, Dan Chirtoaca, and Martin Avagyan. *As We May Study : Towards the Web as a Personalized Language Textbook*. 2018. DOI: [10.1145/3173574.3173912](https://doi.org/10.1145/3173574.3173912).
- [11] Shadan Malik. *Enterprise Dashboards*. Vol. 1. 2005, p. 243. ISBN: 9788578110796. DOI: [10.1017/CB09781107415324.004](https://doi.org/10.1017/CB09781107415324.004). arXiv: [arXiv:1011.1669v3](https://arxiv.org/abs/1011.1669v3).
- [12] *Misunderstood Metrics: Time on Page / Average Session Duration | Analytics Edge Help*. URL: <https://help.analyticsedge.com/article/misunderstood-metrics-time-on-page-session-duration/> (visited on 06/08/2018).
- [13] Yeonjeong Park and Il-hyun Jo. "Development of the learning analytics dashboard to support students' learning performance." In: *Journal of Universal Computer Science* 21.1 (2015), pp. 110–133. ISSN: 09486968. DOI: [10.3217/jucs-021-01-0110](https://doi.org/10.3217/jucs-021-01-0110). arXiv: [110-133](https://arxiv.org/abs/110-133). URL: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84933039245&partnerID=tZ0tx3y1>.
- [14] Eike Pierstorff. *How do google analytics detect the user leaves page? [Online forum comment]*. 2014. URL: <https://stackoverflow.com/questions/27573879/how-do-google-analytics-detect-the-user-leaves-page/27573995#27573995>.
- [15] David D. Qian. "Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension." In: *Canadian Modern Language Review* 56.2 (1999), pp. 282–308. ISSN: 0008-4506. DOI: [10.3138/cmlr.56.2.282](https://doi.org/10.3138/cmlr.56.2.282). URL: <http://www.utpjournals.press/doi/10.3138/cmlr.56.2.282>.
- [16] Cristóbal Romero, Sebastián Ventura, and Enrique García. "Data mining in course management systems: Moodle case study and tutorial." In: *Computers & Education* 51.1 (2008), pp. 368–384. ISSN: 0360-1315. DOI: [10.1016/J.COMPEDU.2007.05.016](https://doi.org/10.1016/J.COMPEDU.2007.05.016). URL: <https://www.sciencedirect.com/science/article/pii/S0360131507000590>.
- [17] Jose Luis Santos, Katrien Verbert, and Erik Duval. "Empowering students to reflect on their activity with stepup!: Two case studies with engineering students." In: *CEUR Workshop Proceedings*. Vol. 931. 2012, pp. 73–86.
- [18] Alexandru C. Telea. *Visual Analytics for Big Data: Application Design*. 2017. URL: <http://www.cs.rug.nl/svcg/VisualAnalytics/Slides>.
- [19] *Understanding Google Analytics 'Time on Page' and 'Session Duration' | Branded3*. URL: <https://www.branded3.com/blog/understanding-google-analytics-time-page-session-duration/> (visited on 06/08/2018).

- [20] Katrien Verbert, Erik Duval, Joris Klerkx, Sten Govaerts, and José Luis Santos. "Learning Analytics Dashboard Applications." In: *American Behavioral Scientist* 57.10 (2013), pp. 1500–1509. ISSN: 0002-7642. DOI: [10.1177/0002764213479363](https://doi.org/10.1177/0002764213479363). URL: <http://journals.sagepub.com/doi/10.1177/0002764213479363>.
- [21] William R. Watson and Sunnie Lee Watson. *An argument for clarity: What are learning management systems, what are they not, and what should they become?* 2007. DOI: [10.1007/s11528-007-0023-y](https://doi.org/10.1007/s11528-007-0023-y). arXiv: [hal-00692067](https://arxiv.org/abs/hal-00692067).
- [22] Stuart Webb. "The effects of repetition on vocabulary knowledge." In: *Applied Linguistics* 28.1 (2007), pp. 46–65. ISSN: 01426001. DOI: [10.1093/applin/aml048](https://doi.org/10.1093/applin/aml048).
- [23] Wikipedia. "Activity tracker." In: *Wikipedia* (2016), pp. 1–2. URL: https://en.wikipedia.org/wiki/Activity_tracker.