SESSION TRACKING FOR AN INTELLIGENT TUTORING SYSTEM

CARLOS HUMBERTO PAZ RODRIGUEZ
Msc. in Computing Science

Supervisors:
Dr. Mircea Filip Lungu
prof. Dr. Alexandru C. Telea

Learning analytics

Software Engineering Department
Faculty of Science and Engineering

June 2018

ABSTRACT

A recurrent problem in online learning is that teachers lack the tools to supervise and evaluate the work done by their students. Accurate and fine grained activity tracking increases reliability on the data analysis and enables a more transparent monitoring of the learning progress. A new activity tracking algorithm is proposed as a mean to provide reliable information about students' activity, and implemented in a language learning platform named Zeeguu. The obtained results enable new types of analysis and a better understanding of the student's efforts.

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

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.

Jugo, Kovačić, and Slavuj [8], Romero, Ventura, and García [16] and Dogan and Camurcu [3] have used data mining algorithms to extract relevant patterns in the learning process. Typical approaches are association rules to discover relations between learning sequences or content. Clustering techniques are commonly used to subclassify a group of students into interesting categories and, for example, detect users with learning difficulties. And some more advanced implementations strive 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 known visualizations that are easily interpreted by untrained users, they are commonly used on dashboards and help summarizing information.

Verbert et al. [20] have performed a comprehensive comparison of different analytical dashboards for ITS systems. They have 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). Almost all of the dashboards focus on the sense making of the learning work done so far, but most of them do not follow up the change of behavior in the users after the dashboard usage.

Others studies investigate how to improve the effectiveness of dashboards in learning activities by promoting motivation. Motivation is a key factor in the success of a learning platform, especially due to the fact that there is not direct physical supervision. Gamification [7] has been proposed as a way to tackle lack of interest due to monotony. 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] [13]. 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 3.2.

1.1 MOTIVATION

With the dynamic capabilities of the Web, a personalized textbook for learning foreign languages [11] [10] is an ongoing research project. The tool, named Zeeguu, allows users to read and learn new vocabulary by recommending texts of their own interest.

A pilot test of the system was implemented in 3 high school classrooms in the Netherlands. All of them learning French as a foreign language. Both students and professors showed an interest in using the proposed tool. One of the outcomes from this research was that professors wanted to have monitoring functionalities that would allow them to keep track of their students' work and progress.

Two major concerns are that students can report false information regarding the work they have been doing with the application, and that 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 accurately can time spent learning a foreign language be measured in an online learning platform?

The results of this research are applicable to the rest of ITS as well as any system that needs time tracking functionality.

1.2 OUTLINE

In Chapter 2, 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 described

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 Chapter 3, 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 explained while implementing it for the reading and exercise activities.

In order to find the best implementation settings, Chapter 4 describes the different tests that were run to find the optimal settings for both implementations. based on the collected data, the best parameter for the reading and exercise timeouts were defined.

Chapter 5 provides an exploratory analysis of students activity after the implementation of session tracking. Visualizations that show how much time, when and how students use the system are presented. Comparisons between students and over time are enabled discussed as a way to provide self awareness.

A conclusion of the findings and difficulties of the proposed solution are described in Chapter 6.

Finally, future research ideas are outlined in Chapter 7.

2 RELATED WORK

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/) [14] [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].

Previous studies focus on analyzing grades and overall performance for students in different courses and within a compared them with the rest of the class. While the "personal" nature of Zeeguu, implies that the information should be tracked for a specific article in a particular language.

Personalized Textbook [11] is a didactic solution that uses personal preferences to provide interesting reading texts and exercises based on those texts. This implies that:

- the learning material is not the same for every student,
- the translated words are individual,
- the practiced vocabulary is completely unique and
- the language(s) being learned can differ.

Therefore:

- The information collected for one user might not be the same as for the rest.
- The learning road is different for each student.
- It is more difficult to compare a student's learning performance against others.

This project focuses on tracking each individual reading and practice session, to provide a detailed usage information. In consequence, a new method for collecting and analyzing learning time is required.

2.1 ACTIVITY TRACKING

The term activity tracking has recently emerged and is mostly related to wearable devices that monitor fitness data [22]. In a more abstract sense, activity tracking is **the process of collecting user's information over time for a particular purpose**.

On the web, technologies such as Google analytics make use of Javascript listeners to 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 [15].

Even wearable devices [6] which are physical dedicated machines, are not infallible. Under certain external condition they can make wrong measurements, therefore 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 [17]. 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 process 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. Automated tools such as RescueTime, track time spent on an application or on a website. But they 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 invested learning a foreign language.

2.2 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 [14].

A dashboard provides timely and useful information so that a decision can be made consequently. Verbert et al. describe the mental process that the user of a dashboard performs to make sense out of it [20]. 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 [18]. 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) [12]

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.

3 METHODOLOGY

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 Dutch, 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 section 3.1, 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 section 3.2.

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

3.1 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.

The main conclusions of these interviews are that teachers would like to:

1. Know how much time their students are investing in the foreign language
2. Have a mechanism to avoid cheating
3. Understand how focused students are when practicing

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

3.2 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, while an exercise session 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.

3.2.1 *Algorithm*

Working sessions are computed under the assumption that all user actions fall into one out of three types: opening, interaction or closing events.

Depending on the type of event, a working session can be either created, updated (kept alive) or finalized. With this simple approach, the algorithm can be fine tuned by adding or removing events (e. g., including the scrolling as a interaction event).

Additionally, there is a **session_timeout** parameter which is used to handle scenarios where a session was not properly closed. In these cases, 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.

Typical scenarios are:

- When the browser was closed or the computer was turned off.
- When the user is inactive for a long period.
- When the user is taking unusually too long between actions the activity is considered as suspicious.

Figure 3.1 depicts the life span of a working session. The circles represent the states in which a session can be:

- Black circle: this is the initial status, this means there is no active session for the user within the current context (I. e. no active reading session for the current article and user, or not active exercise session for the user).

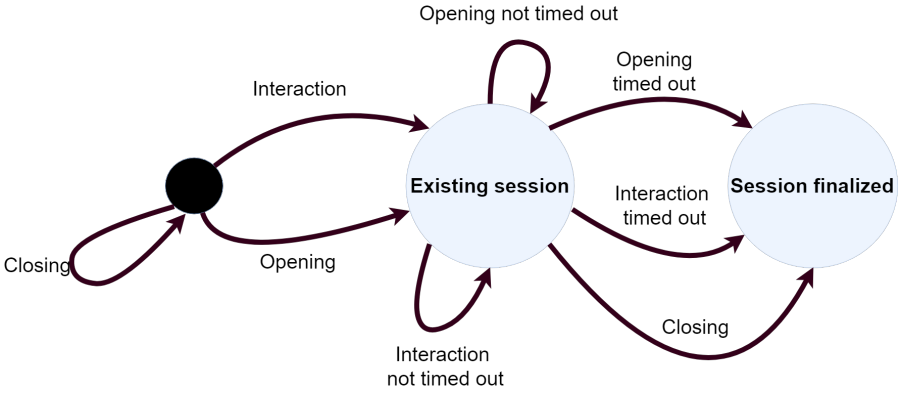


Figure 3.1: Finite state diagram - Working session transitions due to user actions

- Existing session: if there is already an existing/active session in the same context, the session falls in the second status.
- Session finalized: the moment that the session is closed due to a timeout or the user performing a closing action. After this status, no further actions can modify the session.

The arrows represent the type of user actions that trigger the change of status of the session:

- Opening: actions that mark the beginning of a working session, such as beginning to read or to exercise.
- Opening timed out: opening actions that happen after the timeout has expired, this can lead to a different status.
- Opening not timed out: opening actions that happen before the timeout is expired.
- Interaction: most of user actions fall within this category, they are the actions that imply the student is actively working.
- Interaction timed out: interaction actions that happen after the timeout is expired
- Interaction not timed out: interaction actions that occur before the timeout is reached
- Closing: when the user manually finalizes the current activity.

3.2.2 *Architecture*

For the specific case of Zeeguu platform, the implementation of the algorithm is divided in two layers: front and back ends. The front end is implemented in Javascript and it tracks the user actions with

the system. The back end is implemented in Python and is where the logic for the working sessions is implemented. The description and link to the actual implementation of the session tracking for the Zeeguu system, is located in Table 3.1.

FILE NAME	ENTRY POINT	DESCRIPTION
user_reading_session.py	update_reading_session()	Python class implementation of the reading session tracking. Contains the logic regarding when to create, update or close a reading session.
user_exercise_session.py	update_exercise_session()	Python class implementation to handle exercise session tracking. Contains the logic regarding when to create, update or close an exercise session.

Table 3.1: Python code with the implementation of the session tracking for the Zeeguu platform.

Figure 3.2 shows the architecture of the solution. The implementation is divided in two layers, which can work independently.

Relevant user actions are detected by the front end layer of the system, and stored in the database. The back end periodically retrieves the list of user events and stitches them (see link to GitHub code in Table 3.1) to extend the session or creates a new session depending on the type of event and the time between the actions. Therefore inserting the session information back in the database, making it available to be displayed. The front end layer can also trigger the execution of the session stitching module after each user activity, enabling real time session analysis. Event stitching is the process of creating or expanding a session between two consecutive user actions.

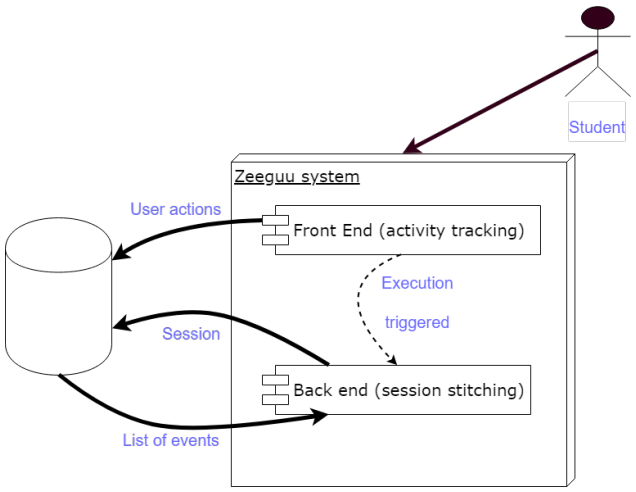


Figure 3.2: Session tracking architecture

3.2.3 *Reading session*

For the reading session, Javascript listeners are used to track the following user events:

- Open list of articles
- Open an article
- Translate a word
- Undo a word translation
- Lose and gain focus on the page
- Scrolling
- Close an article

When these events are detected, the related information is sent to the server and stored in the database. Each time the back end runs, it stitches the events to the corresponding session, until the article is closed or a new article is opened. In which case, the back end finalizes the session, making it available to be visualized.

While defining what a session is in the system, and given the fact that the application is on the web, the following difficulties are encountered:

- Events like closing the browser or turning off the computer cannot be detected
- If the user walks away from the screen, we cannot detect exactly when it was and if he read a couple of minutes before leaving
- The user can open multiple instances of Zeeguu, and each instance can have a different article open
- If the user opened multipe articles, the user can change from article to article in a short time period
- The user can be using more than one device at the same time
- The user can read an article in multiple sessions

Therefore the following assumptions for the implementation of the reading session are considered:

- Reading sessions are considered per article. Therefore, when the user opens a new article we consider it as a new session. In this way we solve the issue where the same user could have opened different articles in different tabs 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.
- A user can read the same article multiple times, and each time is considered a new session.

As mentioned before, a timeout parameter is used for finalizing a 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. Therefore, defining the value of the timeout is crucial.

3.2.4 *Exercise session*

For the exercise sessions, the front end layer starts a timer when the exercise is opened. When it is answered (either correct or incorrect) or a hint is requested, the event is stored in the database with the elapsed time.

The back end layer is immediately executed after the front end code, appending the recently answered exercise to the exercise session or creating a new one when the elapsed time is larger than the timeout value.

Therefore, opening an exercise does not count for the active session, only answering.

For exercise sessions, the algorithm 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.

4 RESULTS

In this chapter, experiments are performed to find the best implementation/settings of the algorithm. As explained in section 3.2.3, a wrong value of the **timeout** parameter can produce different and perhaps incorrect results. First the session timeout for the reading sessions will be determined by trying and analyzing different parameter values. Second, a the session timeout for the exercise sessions will be determined.

4.1 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.

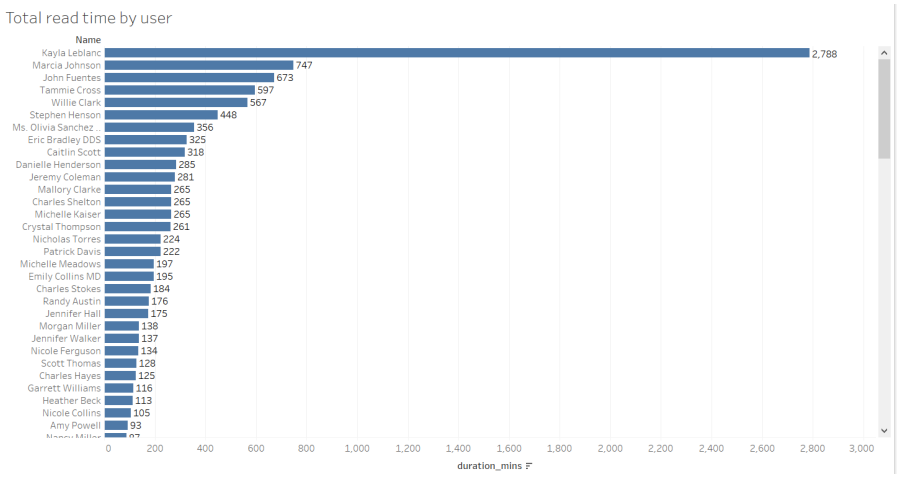
User actions were classified as follows:

- Opening
  - Article focused
  - Open article
  - Open starred article
- Interaction
  - Change orientation
  - Open alter menu
  - Close alter menu
  - Send suggestion
  - Speak text
  - Translate text
  - Undo text translation
  - Scroll
- Closing
  - Article closed
  - Articles requested from Zeeguu

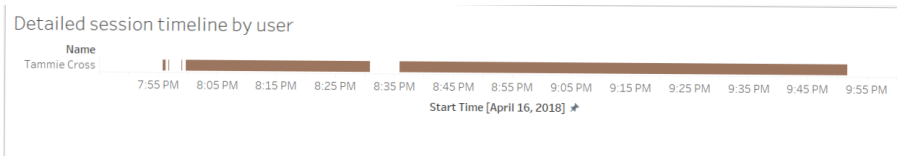
4.1.1 *Baseline*

For the baseline an arbitary value of 5 minutes for the **session_timeout** was used.

By observing the total reading time by user, we can see that the highest value is 2788 minutes (almost 2 days). By analyzing the daily activity for a specific user, we find long reading sessions spanning up to 1:20 hours long. Figure 4.1 show these results.



(a) Total read time by user



(b) Daily activity timeline

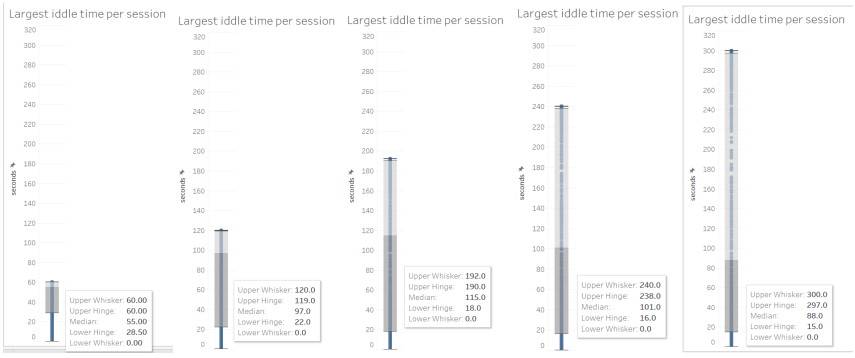
Figure 4.1: Reading session visualizations for session_timeout of 5 minutes.

4.1.2 *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 grace time for a user reading without doing any action. Scrolling events are helpful to detect more frequent activity, but they can be overwhelming for the web server and the database, for that reason, the web server callback was limited to maximum 1 call per minute, for the scrolling event. This implies that the minimum timeout value should be at least 1 minute.

Sessions were computed using different parameters for the session_timeout, and the maximum idle time between user actions inside a session was plotted (Figure 4.2).

We can observe that the median value fluctuates between 60 and 120 seconds (as shown in Table 4.1). By taking this value into account



(a) 1 min (b) 2 min (c) 3 min (d) 4 min (e) 5 min

Figure 4.2: Box plot of maximum idle time per session, for distinct session timeout values

and some empirical measures, the value of the timeout was set to 2 minutes.

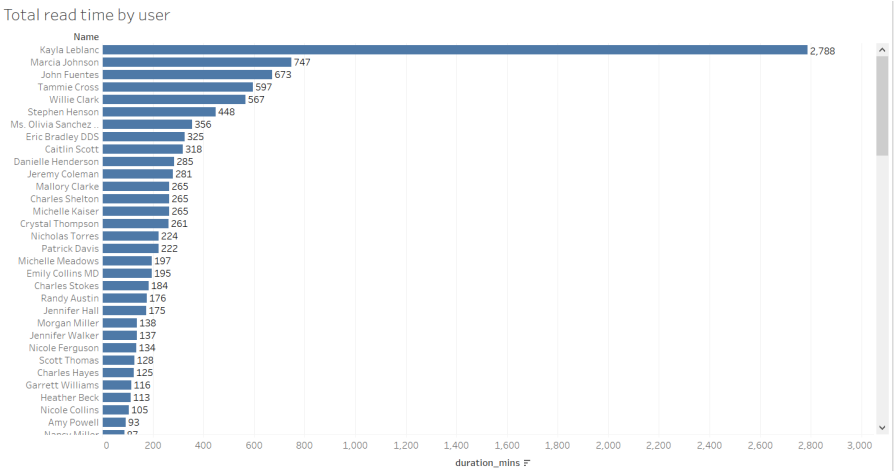
By comparing the medians of the tests, we can observe that the maximum idle time for the tested timeout values is located between 55 seconds and 115 seconds. If we would chose a timeout value of lower than 115 seconds, we would be closing sessions while the student would still be working. Therefore a value of at least 115 seconds should be used. Given that the timeout value is a continuous variable and that we are only testing 5 possible settings, we decided to be conservative and round up the timeout value to 2 minutes.

The upper hinge, in contrast with the median, increases as the timeout value do. This is because small sessions that are close enough, are stitched together, therefore moving the maximum value upwards. In other words, some noise is introduced in the analysis.

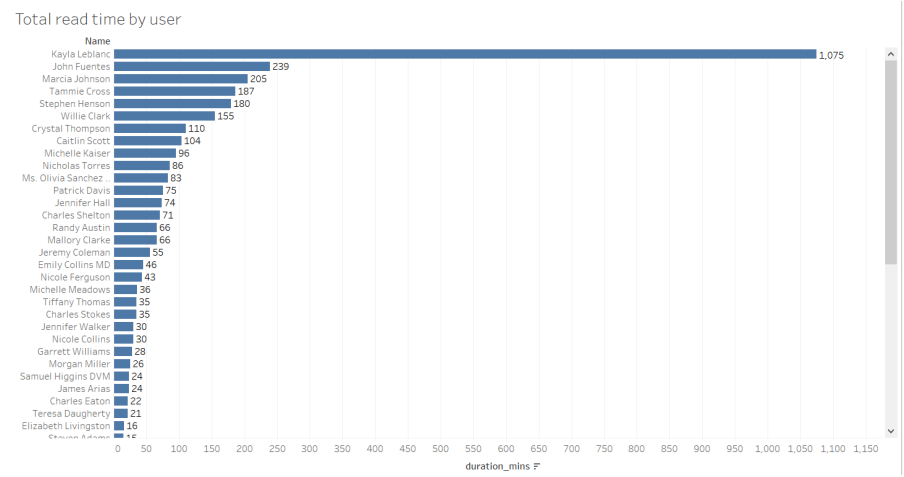
TIMEOUT VALUE (MIN)	MEDIAN (SEC)	UPPER HINGE(SEC)
1	55	60
2	97	119
3.2	115	190
4	101	238
5	88	297

Table 4.1: Maximum idle time comparison between different session timeout values

By comparing the total time per user between the baseline and the final settings (Figure 4.3), we can observe that a reduction of almost 300% in the total time per user was achieved. This means that the session_timeout value has a huge impact, specially when comparing over long periods of time.



(a) Baseline (5 minutes)



(b) Final settings (2 minutes)

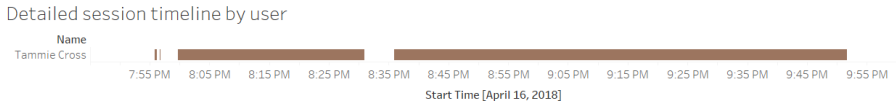
Figure 4.3: Total time improvement comparison

Finally, for the detailed activity, a big change is observed, in Figure 4.4a, we observe roughly two smooth reading sessions, which later, with a finer session detection gets split into smaller reading sessions (Figure 4.4b). This suggests that we correctly detected the user's gap of attentions, therefore a more precise session time was measured.

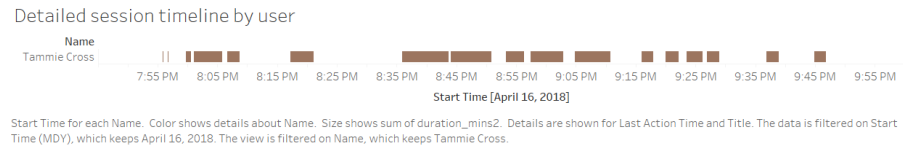
4.1.3 *Validation*

Before the implementation of the session tracking, students manually reported the time spent reading in Zeeguu. Therefore a useful validation of the computed session tracking is against the time reported by students.

Table 4.2 shows the comparison between the reported time and the computed total time using session tracking for three students over a period of six months.



(a) Baseline (5 minutes)



(b) Final settings (2 minutes)

Figure 4.4: Low level session tracking time line

There is a clear difference between both times. The reported time is 4 times bigger than the computed time.

The validation against the time reported by students did not match as expected, a possible reason is that students rounded up to the next half hour while generating the reports because that is their unit of measure, according to the teacher interview (Interview A).

STUDENT	READ TIME	EXERCISE TIME	TOTAL DETECTED TIME	REPORTED TIME
Student 1	59 min	1 hour 24 min	2 hours 23 min	10 hours
Student 2	44 min	0 min	44 min	2 hour 30 min
Student 3	54 min	2.8 min	56.8 min	4 hours

Table 4.2: Comparison of working time reported by students vs the time computed by the algorithm (using a timeout value of 2 minutes for reading sessions and 21 seconds for exercise sessions)

4.2 EXERCISE SESSION EXPERIMENTS

For the exercise sessions, the implementation is simple. Exercise user events were classified as follows:

- Opening
  - Opening exercise
- Interaction
  - Answer exercise
  - Asking for a hint

Given that answering marks the end of an exercise, no additional "closing" actions were used. Therefore the exercise session is finalized whenever the timeout expires. However, in contrast with the reading session, for exercise sessions no grace time is given, meaning that

the time of the end of the session is marked by the last completed exercise, and not when the timeout expires.

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.

In this case, the time between the exercise opening and the answering is computed by a timer. Therefore by utilizing a descriptive analysis of the answering time (box plot in Figure 4.5). The obtained results are 6, 10 and 21 seconds, which represent the median, upper hinge and upper whisker. These metrics correspond the 50%, 75% and 100%, of the sessions at hand.

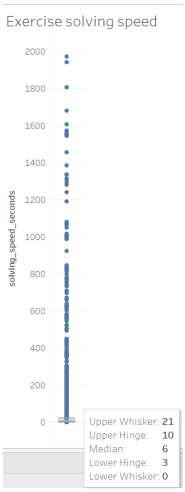


Figure 4.5: Exercise solving speed boxplot

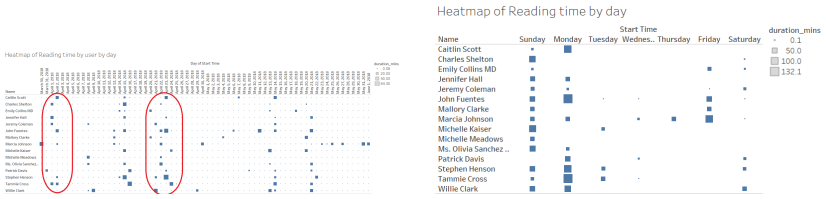
Given that the tracked time window is so small, a final setup of no grace time but using the upper whisker value (21 seconds) is chosen.

5 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 chapter. With this knowledge they can adjust their teaching strategy and have a better understanding of how the class is using the system.

5.1 READING SESSION

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 5.1a). These can be explained with a visualization by week day. In Figure 5.1b we can confirm that the students are mostly actively reading on Sundays and Mondays.



(a) Heat map of time spent by day and user in a class

(b) Heat map of time spent by weekday and user in a class

Figure 5.1: 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 5.2 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.

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 5.3, an example of a dashboard containing the information that could be relevant for a user is displayed.

5.2 EXERCISE SESSION

Typical questions related to the exercises are how much time has the student been practicing? Do they practice daily? For how long? Etc.

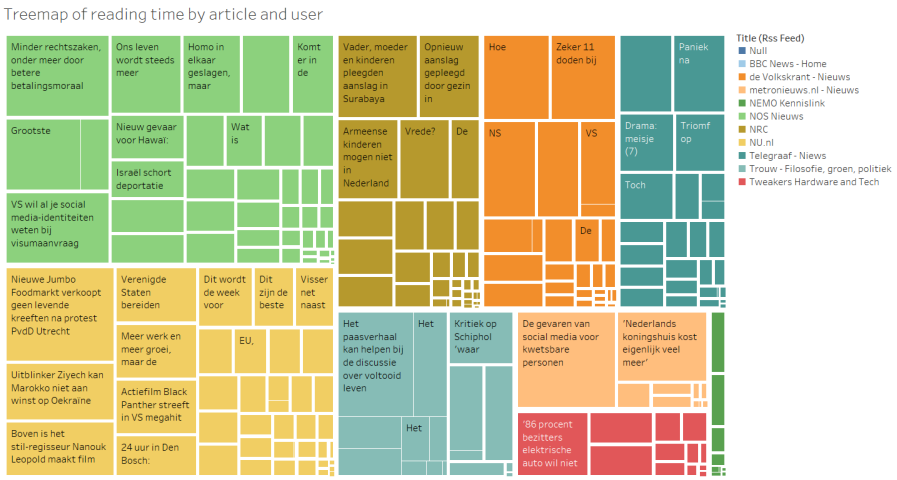
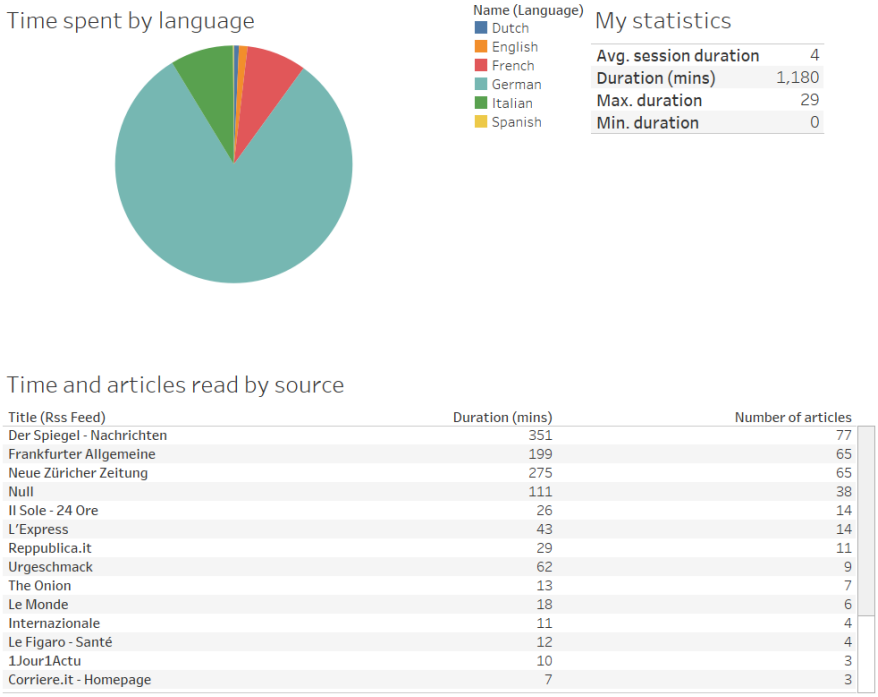


Figure 5.2: Treemap of time spent by article. The color indicates the source of the article. Each block represents a user and the size represents the total time spent by a user on that article



Title (Rss Feed)	Duration (mins)	Number of articles
Der Spiegel - Nachrichten	351	77
Frankfurter Allgemeine	199	65
Neue Züricher Zeitung	275	65
Null	111	38
Il Sole - 24 Ore	26	14
L'Express	43	14
Reppublica.it	29	11
Urgeschmack	62	9
The Onion	13	7
Le Monde	18	6
Internazionale	11	4
Le Figaro - Santé	12	4
1Jour1Actu	10	3
Corriere.it - Homepage	7	3

Figure 5.3: Dashboard displaying individual statistics about reading habbits

In this chapter 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 5.4, some students have been constant in the amount of time

spent practicing between 2017 and 2018, while others have decreased or even stopped it completely.

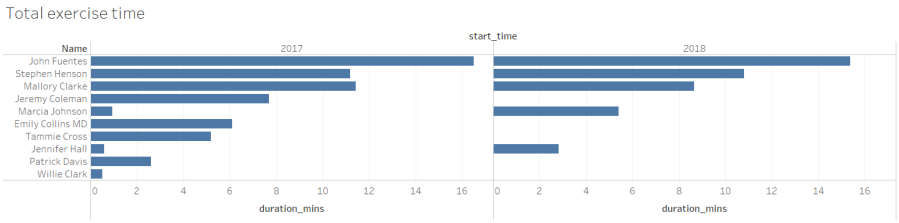


Figure 5.4: Class exercise time year over year comparison

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 5.5).

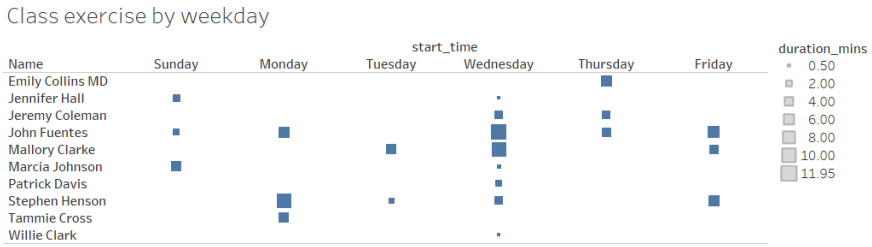


Figure 5.5: Class exercise heatmap by weekday

For understanding the quality of the practice, we can also visualize the average session length. An interesting finding, as we can see in Figure 5.6, 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 or about their knowledge, since exercises have fixed number of questions, and thus, a fast student will finish faster. 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.

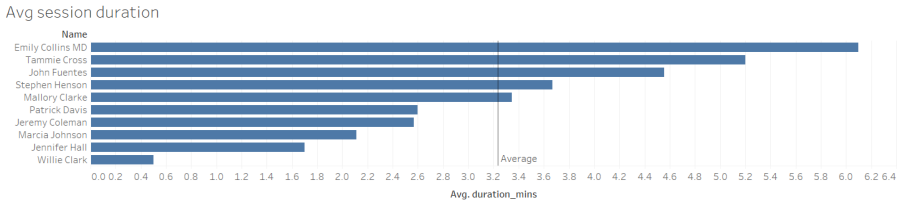


Figure 5.6: Bar chart showing average session duration by student

Continuing with the analysis of the working habits, a detailed time line (Figure 5.7) 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.

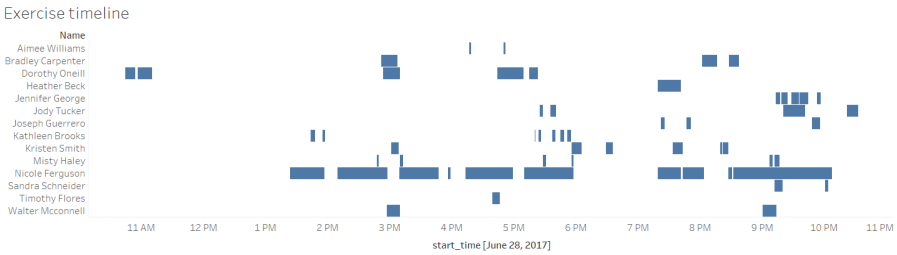


Figure 5.7: Class exercise time line

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 5.8 we observe that the student was not constant in his daily practice.

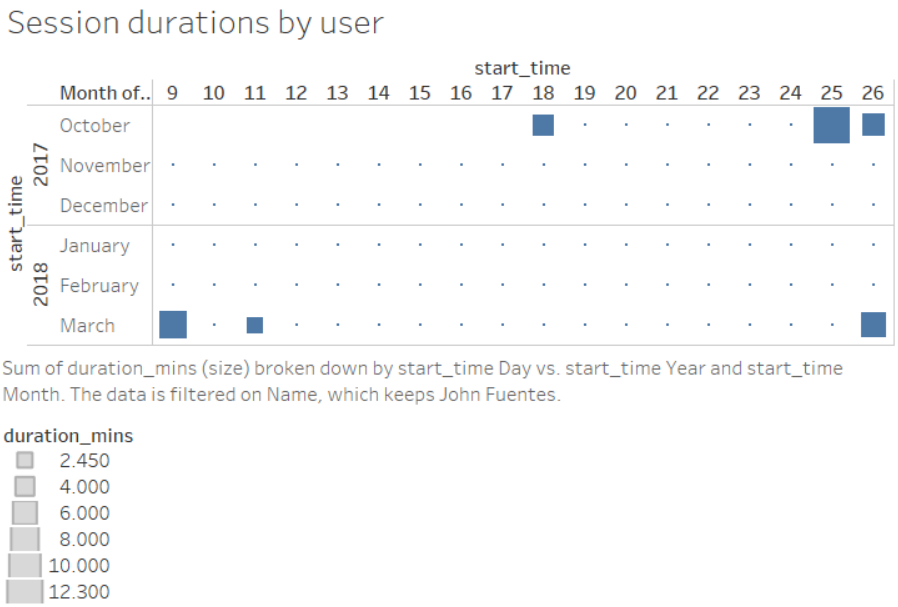


Figure 5.8: Heat map showing exercise time calendar

6

CONCLUSION

As observed in Chapter 5, not only the time spent but also the work patterns and habits can be spotted with a precise session tracking. Students who work uninterruptedly versus those who lost focus very frequently can be easily detected (Figure 5.7). Then, teachers 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. Outside of the learning context, the session tracking algorithm could also 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).

7

FUTURE WORK

One of the main challenges for defining the **session_timeout** is having statistically representative data. In this particular implementation, the Zeeguu system has been recently deployed and has suffered constant modifications (e. g. scrolling event detection was included during the implementation of this project), 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 is 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 might 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 rise an alert for those who are under performing.

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 world 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. arXiv: 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. 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. 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.

[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. 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.

[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. 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 Informa-*

tion and Communication Technology, Electronics and Microelectronics, MIPRO 2014 - Proceedings. 2014. ISBN: 9789532330816. DOI: 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.

[10] Mircea F. Lungu. "Bootstrapping an Ubiquitous Monitoring Ecosystem for Accelerating Vocabulary Acquisition." In: *Proccedings of the 10th European Conference on Software Architecture Workshops*. ECSAW '16. Copenhagen, Denmark: ACM, 2016, 28:1–28:4. ISBN: 978-1-4503-4781-5. DOI: 10.1145/2993412.3003389. URL: http://doi.acm.org/10.1145/2993412.3003389.

[11] 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.

[12] Shadan Malik. *Enterprise Dashboards*. Vol. 1. 2005, p. 243. ISBN: 9788578110796. DOI: 10.1017/CBO9781107415324.004. arXiv: arXiv:1011.1669v3.

[13] *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).

[14] 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. arXiv: 110-133. URL: http://www.scopus.com/inward/record.url?eid=2-s2.0-84933039245{\&}partnerID=tZOtx3y1.

[15] 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 (visited on 06/06/2018).

[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. 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. 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. arXiv: hal-00692067.

[22] Wikipedia. *Activity tracker*. 2016. URL: https://en.wikipedia.org/wiki/Activity{_}tracker (visited on 06/07/2018).