TraceMe + Pedagogical Agent = Morgan: Personalized Learning With AI

Fatimah Wirth, Ph.D.

Abstract

Online courses are often structured using a cookie cutter approach and even when they are personalized, they are done based on instructor/instructional designer experience. In this research study, the learners' paths was traced throughout the online course. Georgia Tech Professional Education (GTPE) collaborated with Professor Alain Mille from the French National Center for Scientific Research and France's Universite Numerique and Pierre-Antoine Champin from Claude Bernard Lyon 1 University in France and the LIRIS research center and used their proprietary software called TraceMe to collect data in a Georgia Tech Professional Education (GTPE) course. The research question addressed was how learners navigate through an online course. The findings from this research study will help identify trends/paths that can be used to restructure curriculum and presentation of content. Understanding the learner's navigational and interactivity patterns will inform effective course design (Stein, 2014). The findings will also help create an artificial intelligent pedagogical agent that provides a personalized path. The AI pedagogical agent will pop up at the point where the learner is frustrated or stuck.

Introduction

The Covid-19 pandemic has forced everyone to learn in new ways. It has ripped us from the traditional classroom environment and forced us to think of education and learning in new and novel ways which may have been unthinkable pre-Covid-19.

Although this research study predates the Covid-19 pandemic, it even more relevant now. Once the research data had been analyzed, the Georgia Tech Professional Education (GTPE) researchers will be able to determine how the learners navigate through a course based on the trace data collected. This will help identify trends/paths that can be used to restructure curriculum and presentation of content. Understanding the learner's navigational and interactivity patterns will inform effective course design (Stein, 2014). Georgia Tech Professional Education collaborated with Professor Alain Mille from the French National Center for Scientific Research and France's Universite Numerique and Pierre-Antoine Champin from Claude Bernard Lyon 1 University in France and the LIRIS research center and used their proprietary software called TraceMe to collect data in a GTPE course.

The aim of this research project was to trace the path that a learner takes during an online course session using javascript applied to each course page within the Learning Management System. The path that is traced includes how the student peruses the course menu and the content within a course. The data collected from the trace of the learner's path through the online course can be beneficial in that the researchers will be able to find out where learners have issues within the course. The issues may be due to navigation, instructions or course content. As a result, the researchers will be able to discover trends or themes when the data is coded. These trends or themes will lead to better online course structure, course content and course navigation. This will

then improve the quality of the online courses so that participants get a more positive and beneficial online course experience.

Research Questions

The aim of this research project is to trace the path that a learner takes during an online course session using javascript applied to each course page within the Moodle Learning Management System (LMS). The path that is traced includes how the student peruses the course menu and the content within a course. This will allow us to identify areas within the course that need improvement.

The research questions investigated in this research study were as follows:
- Identify where and when learner engagement drops off
- Identify where content needs to be clearer or re-emphasized

An IRB protocol was submitted and approved before research and data collection started.

Literature Review

New challenges in information technology has forced institutions of higher learning to rethink education. Many institutions have turned to Artificial Intelligence (AI) to improve education. According to Yolvi, et al.:

> The formats based on artificial intelligence promise a very substantial improvement in education for all the different levels, with an unprecedented qualitative improvement: providing the student with an accurate personalization of their learning tailored to their requirements, managing to integrate the various forms of interaction human and information and communication technologies.
>
> (Yolvi, et al., 2019, p.536)

Artificial Intelligence creates computing systems that have the ability to learn from information in its environment and provide adaptive behaviors to assist learning. For the past 30 years, many institutions of highly learning have focused on creating AI based systems to mimic human tutoring. These systems are highly autonomous, interactive and adaptable (Qin, et al., 2020).

AI can also assist those with intellectual disabilities. These learners need a wide range of learning needs and do not learn in a linear or hierarchical way. The use of artificial intelligence tools for education can help move a student from a negative state to a positive state of learning. Students who are frustrated or stuck at a certain place in a course can be nudged to move into a more positive state of engagement. (Standen, et al., 2020)

However, so far, even though there may be intelligent tutors powered by AI, there is yet not a single one that provides just-in-time assistance where the AI pedagogical agent actually pops up at the time of frustration or being stuck in one place in a course and lets the learner know that there are other pathways that they can follow in order to get unstuck. This research study aims to do just that.

Methodology

The instructional designer reached out to at least three instructors teaching Georgia Tech Professional Education online courses in Fall 2017. An instructor was chosen based on the

interest of the professor to participate and the ease of navigation of the course. The duration of the subject participation would be the duration of the course.

A week before the start of the course, the students in the course were sent a Qualtrics survey asking for voluntary participation in the research study. The students had to enter their first name, their last name and their email addresses. A consent form was attached to the email. In addition, before taking the background or end of course survey, the participants in the research study were provided with verbiage that said that if they took the survey, then they were consenting to taking part in the research.

The students first name, last name and their email address were asked so that the sample population would be identified. Once this population has been identified, their actual trace data will be anonymous. Their names will not be attached to the trace data in any way. The system will generate a random ID for each student based on CAS authentication. So each time the student logs into the course, the same random ID will be used for the student. There will not be a key or any document that will link the names to the random ID generated at any point during the study.

A new role called "Non Participant" was created for all those who did not fill out the survey or who did not want to participate in the research study. The new role was manually added to each student who did not want to participate. By adding this role, these students would not be tracked. These students would therefore have two roles, namely the Student role and the Non Participant role. This exclusion from participation in the study was created via the use of an HTML block in the Moodle Learning Management System used by Georgia Tech Professional Education. The HTML block would be called KTBS. The Non Participant role will ensure that the students who do not wish to participate will not see the KTBS HTML block when they logged into their course. This would mean that they will not participate in the study. The students who agreed to participate in the research study will see the KTBS HTML block when they log into their course. In addition, they will see a large red dot on the top right of the page.

When the students who had agreed to participate in the research study logged into the course, their every action in the course would be traced. This was an automatic process and the students would not have to do anything to initiate the process. If the students who had agreed to take part in the study did not want their actions traced, they can click on the red dot. The green dot will then become a red dot. This will mean that their actions in the course are not being tracked. If they log out of the course and then log back in, then the green dot will automatically appear again. This means that they will have to uncheck the check box again if they do not want to be tracked. The tracing of the data will not impact or change the normal user navigation or course setup in terms of activities and assessments.

The students were informed that there were no known risks for participating in the study. All student information would be kept confidential. All records would be kept under a code number and not by student name. However, students were warned that data may exist on backups or server logs after the end of the research project. All data would be deleted three years after the completion of the study.

The students were also informed that the trace data would be on the Amazon Cloud server. The data collected would be anonymous and only the researchers would have access to this server. Only the researchers would be able to download the data as a CSV file.

The students who agreed to participate in the course were identified through the response to the survey and given access to a consent form. In the consent form, students were notified that the software called "TraceMe" would be activated when they logged into the course. They would

see a red dot next to their names on the top right and that signified that their actions in the course was being traced (see image below).
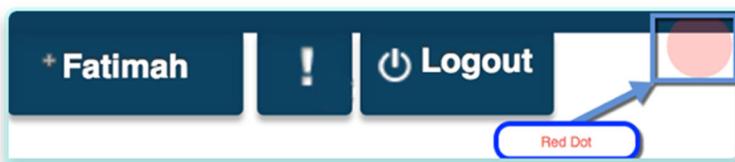
Red Dot

These students were also informed that if they wished not to be traced on a certain day or if they no longer wished to have their actions in the course traced, then they could click on the red dot next to their names on the top right. The red dot would then turn into a grey square and this would denote that they were no longer being traced (see image below).

Grey Square

The students who did not wish to participate would not have access to the software and hence would not see a red dot next to their names on the top right. They would participate in the course as any other student in an online course would.

In addition, the participants will have to fill out a background survey and an end of course survey. The background survey will collect demographic information of the participants. The end of course survey will ask questions about the course design and navigation.

The background data will be collected once at the beginning of the course through the background survey. The end of course survey data will be collected once at the end of the course through the end of course survey. The end of course survey will determine the difficulty level of the content and the navigational ease of the course. The data collected from the background and end of course surveys will be correlated to the trace data. For example, the trace data and the background data can be correlated to show how younger participants navigate versus how older participants navigate through the course. There may also be other correlations in terms of navigation through the course such as gender, educational background, reason for taking the course, etc. The trace will be continuously sending data on student navigation of the online course to the server. This data will be captured for the entire duration of the course. The data collected will assist the Georgia Tech Professional Education instructional designers to figure out where students get stuck within a course and what can be done to alleviate that. There could be follow up interviews with each student who had difficulties to find out why they got stuck and what would have made either the navigation or the content better meet their needs.

Course Design

The Learning Management System at Georgia Tech Professional Education at the time of the research was Moodle hosted by Moodlerooms. The course chosen for the research project was VET^2: Military Transitions to the Workforce. The course was chosen based on the fact that it had a clean interface and was easy to navigate. The course was also designed based on the

Standards from the Quality Matters Higher Education Rubric, Sixth Edition and had enough activities to make the course engaging.

The tracing software that we piloted was developed by Dr. Alain Mille from Liris (Trace-Based System - a project of the TWEAK Team) and Pierre-Antoine Champin from Claude Bernard Lyon 1 University in France. The trace engine was installed on a server and javascript code was appended to each course page in Moodle. The trace would be automatically activated as soon as the learners who had agreed to participate in the research study logged into their course. The tracing software then would record and archive all paths taken by the learners. The tracing of the data will not impact or change the normal user navigation or course setup in terms of activities and assessments.

The data collected from the trace of the learner's path through the online course can be beneficial in that the researchers will be able to find out where learners have issues within the course. The issues may be due to navigation, instructions or course content. As a result, the researchers will be able to discover trends or themes when the data is coded. These trends or themes will lead to better online course structure, course content and course navigation. This will then improve the quality of the online courses so that participants get a more positive and beneficial online course experience.

Data Collection

Data was collected from the background survey, the end of course survey and the trace from the TraceMe software. Data from the background survey and the end of course survey will be analyzed quantitatively and/or qualitatively. These data may also be exported to qualitative and/or quantitative software packages.

Dr. Champin created a Pilot HTML block in the course, to host the script that would collect the data when the student who had agreed to participate in the research project logged into the course. The student would see the red dot on the top right of the page when data was being collected.

The trace data collected from the Amazon Cloud server would be analyzed qualitatively and/or quantitatively. The data is anonymous and only the researchers have access to this server. In addition, the trace data would be exported to qualitative and/or quantitative software packages. Open coding would be used to discover trends, themes or confounding variables within the qualitative data collected. The data may also be cross-tabbed across multiple data sources (e.g. background survey, end of course survey and trace data from plugin) in order to discover trends, themes or confounding variables. The tracing of the data will not impact or change the normal user navigation or course setup in terms of activities and assessments.

The data collected from the participants of the research study will be kept confidential to the extent allowed by law. Participant records will be kept under a code number and not by name. Participant records will be kept in a locked file if printed and on a secure desktop or laptop if electronic. The data on the desktop or laptop will be kept secure by locking the desktop or laptop screen with a user login and password. When the desktop or laptop computer is not in use the screen will be locked. Only the researchers will have access to these files - printed and/or electronic. Participants in the research study will only be identified by number. Data correlation will use this number to match data from various sources to the participant. Researchers will share data and collaborate on a private Office 365 site. The site will only be available to the

researchers. All data (printed and electronic) will be destroyed three years after the completion of the study.
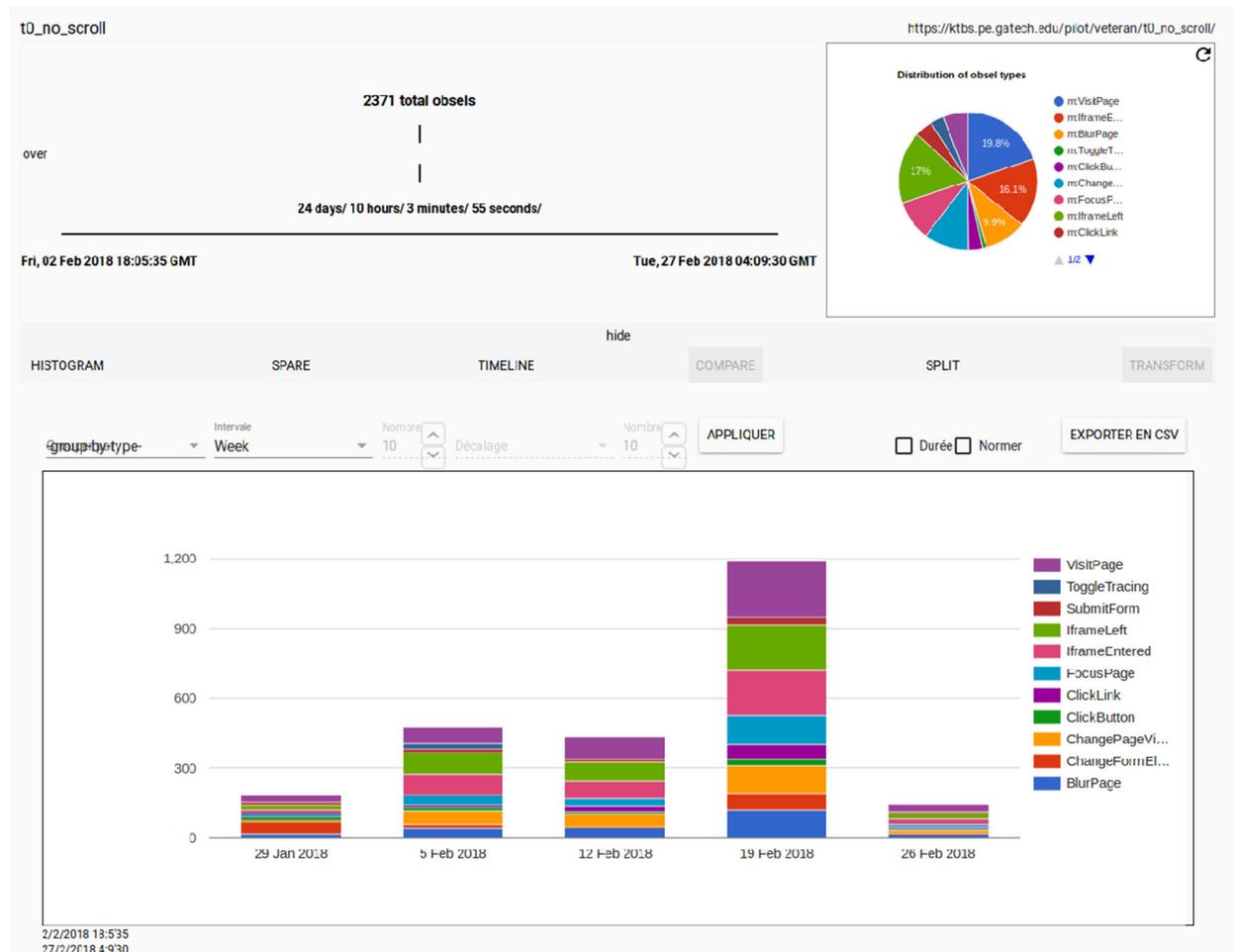
Data Analysis

Data collection took place in Spring of 2018. The chosen course, VET[2]: Military Transitions to the Workforce, went live on February 5, 2018. There were 26 students in the chosen course and all 26 students agreed to take part in the research study. The students were given 6 weeks to complete the course which had 4 modules, 4 end of module quizzes, 5 assignments and 7 videos to watch. Data was collected on an AWS server. The Amazon Cloud external server only collected anonymous data. The researchers had access to this server and were able to download this data as a CSV file.

In just 24 days, 2371 pieces of data was collected. The data collected included visiting a page, submitting an assignment, clicking a link, clicking a button, changing page view.

Figure 1
Preliminary data obtained from VET[2]: Military Transitions to the Workforce course.

Of more interest however, was how the students were making their way through the course and where they got stuck or spent the most amount of time. However, data collection and data analysis had to stop before the GDPR rules went into effect on May 28, 2018 as the researcher collecting the data was Pierre Anton Champin who resided in France. Pierre handed over the data to the researcher in the United States and data analysis has stalled since then because of the new GDPR rules.

Work in Progress

After data collection ceased due to GDPR restrictions, data analysis also ceased. This is due to the fact that the researchers who would be analyzing the data were from France. Currently there are more than 40,000 bits of data waiting to be analyzed.

The aim of the project has also changed from identifying two research problems to the creation of an AI pedagogical agent that would address the two research problems. The first research problem was to identify where and when the learner engagement drops off. The second research problem was to identify where content needs to be clearer or reemphasized. The pedagogical agent, named Morgan, would pop up at the point when engagement drops off. The pedagogical agent will provide either encouragement, explanation or a new path for the learner to follow so that the learner can complete the task. In addition, research partners will need to be identified to assist in the process of the AI creation based on the current data collected.

However, since Georgia Tech Professional Education now uses Canvas as the Learning Management platform instead of Moodle, replication of research needs to be conducted in the same course in Canvas.

References

Arora, A. (2020). Disrupting clinical education: Using artificial intelligence to create training material. *Clinical Teacher*, *17*(4), 357–359. https://doi.org/10.1111/tct.13177

Charlesworth, A. (2014). The comprehensibility theorem and the foundations of artificial Intelligence. *Minds & Machines*, *24*(4), 439–476. https://doi.org/10.1007/s11023-014-9349-3

Heinimaki, T. J., & Vanhatupa, J.-M. (2013). Implementing artificial intelligence: a generic approach with software support. *Proceedings of the Estonian Academy of Sciences*, *62*(1), 27–38. https://doi.org/10.3176/proc.2013.1.04

Morris, Sean M. and Stein, Jared (September 4, 2014). Lossless Learning: An Interview with Jared Stein, Hybrid Pedagogy, Retrieved on 1/15/2015 from https://hybridpedagogy.org/lossless-learning-interview-jared-stein/

Ocaña-Fernández, Y., Valenzuela-Fernández, LA, & Garro-Aburto, L. (2019). Artificial intelligence and its implications in higher education. *Purposes and Representations, 7* (2), 536-552. doi: http: //dx.doi.org/10.20511/pyr2019.v7n2.274

Qin, F., Li, K., & Yan, J. (2020). Understanding user trust in artificial intelligence-based educational systems: Evidence from China. *British Journal of Educational Technology*, *51*(5), 1693–1710. https://doi.org/10.1111/bjet.12994

Standen, P. J., Brown, D. J., Taheri, M., Galvez Trigo, M. J., Boulton, H., Burton, A., Hallewell, M. J., Lathe, J. G., Shopland, N., Blanco Gonzalez, M. A., Kwiatkowska, G. M., Milli, E., Cobello, S., Mazzucato, A., Traversi, M., & Hortal, E. (2020). An evaluation of an

adaptive learning system based on multimodal affect recognition for learners with intellectual disabilities. *British Journal of Educational Technology*, *51*(5), 1748–1765. https://doi.org/10.1111/bjet.13010

Standards from the Quality Matters Higher Education Rubric, Sixth Edition. Quality Matters. Retrieved from Specific Review Standards from the QM Higher Education Rubric, Sixth Edition

Trace-based system – a project of the Tweaked Team. LIRIS. Retrieved 9/23/2020 from https://projet.liris.cnrs.fr/sbt-dev/tbs/doku.php

Wang, S., Yu, H., Hu, X., & Li, J. (2020). Participant or spectator? Comprehending the willingness of faculty to use intelligent tutoring systems in the artificial intelligence era. *British Journal of Educational Technology*, *51*(5), 1657–1673. https://doi.org/10.1111/bjet.12998