Project 326 and Agent Replay
Project Proposal

Prepared for

Derrick Murry/Lily Roark

Capital One

By

Del Huband, Brian Vo, Alex Nguyen, Parker Gonzaga

Under the supervision of

Rachita Sowle

10/10/2024

Executive Summary

Capital One's customer service system Empath has frequent bugs that hinder agents attempting to help customers. The Empath developers currently have a difficult time debugging these issues, as they have to manually read and decipher relevant clickstream data relating to the error. In order to solve this problem, our team will develop a program that can replay an agent's journey while they are on the site so that the developers can easily reproduce bugs that agents may have encountered. This replay could also be used to assist in training new agents, improve agent workflows, or generating automated tests on Empath to test the service's performance or features. However, this program cannot reveal any customer information as to not leak any sensitive information while replaying an agent's journey. It must be scalable to handle multiple agents' sessions, while still maintaining a small file size so it can be handled easier.

To produce this project, we'll be hosting a personal website that will have clickstream events attached to it. We'll read these clickstream events from the website using an analytical tool, and these files will be saved onto a cloud storage so that they can be easily accessed and replayed. Now that the clickstream data is saved, it will be plugged into a browser playback service that'll show an approximation of the user's session in real time, allowing others to view how a user interacted with the website, and potentially find the same issues and bugs that the user encountered.

Table of Contents

Section A. Problem Statement

Capital One is a bank holding company. They provide both in-person and online services. When it comes to online services, customers occasionally encounter issues with their bank accounts, or want to manage their bank accounts. To do this, they contact customer support.

Capital One uses an online service called Empath to manage customer accounts during customer support. Empath is used by customer service agents to assist any customers who need help with their account, such as being locked out of it or needing to change data. The issue arises that any agent using Empath could potentially encounter bugs. This problem occurs semi-regularly, depending on how stable the current version of Empath is. When these issues arise, Empath's developers use clickstream data to track down what caused the issue. This process is time consuming, leading to slower bug fixes and update cycles being deployed for Empath. This worsens customer interactions and hinders employee productivity. Our solution should reduce the amount of time it takes to locate encountered issues, thus improving customer and agent experiences.

Empath deals with sensitive customer details as well as proprietary Capital One information, so using a publicly available solution to solve this problem directly is not the preferred approach. Additionally, Capital One agents do not have the sufficient privileges to use these solutions directly, and the solutions currently available to developers are too time-costly.

The project's client is Capital One. Stakeholders in this project include Capital One agents, developers, training teams, customers, and shareholders. The first three groups can use our project directly, and the time-saving improvements our project produces will affect the experiences of customers and shareholders.

The problem has some associated costs. When agents face a bug that cannot be replicated it can hurt both the customer and the company economically. When an agent isn't able to solve a problem it leads to a loss of productivity for the agent. And for customers if they have an urgent issue that they want to be solved, bugs can delay their business functions, such as a payment.

This project falls under the umbrella of user-interaction tracking, data analytics, and User Experience (UX) research. While leveraging Google Analytics, Heap, Amplitude, or another website analytic tool, a more comprehensive view of user behavior can be generated. This will enhance the accuracy and depth of clickstream analysis for this client, advancing the way that an agent's journey data is tracked, stored, and replayed for optimization.

In 2018, a team composed of Timur Bekmambetov, Guy Bedford, and Daniel Kehoe developed their own service which would playback a user's browser session (Bekmambetov et al., 2018). Their system allowed for users to play and rewind the journey a user took while browsing one or multiple websites using JSON data that tracked the mouse cursor and mouse events, and it could track the user's sessions through multiple websites (Bekmambetov et al., 2018). This JSON data was captured via a program that has to be run by the user to record any

data (Bekmambetov et al., 2018). However, this software was only developed to be used locally, meaning that all computers owned by Capital One would have to have this external software installed (Bekmambetov et al., 2018). It also has no way of mocking client data, and it doesn't automatically record all of a user's sessions on a specific website, meaning that some potentially insightful user sessions could get lost if Capital One were to implement this solution.

One previous solution attempt implemented by Capital One was Glassbox, a software that would record agent actions in videos that could be played back (Glassbox). These screen recordings were fed through a service to censor sensitive data, and then used by developers to see what led up to encountered issues (Glassbox). Some problems with this approach included large storage size for the videos, potential mistakes releasing sensitive customer data, and videos not capturing backend data (Glassbox). Our approach should analyze clickstream data in order to avoid analysis of sensitive data and to illustrate what is happening in the backend.

Another software that is similar is Crazy Egg. It has a lot of features that compared to Glassbox that can be viewed as better. Crazy Egg is more oriented towards providing visual tools, such as heatmaps or snapshots, however it has limits when it comes to the amount of recordings it can produce, which is limited at 10,000 recordings (Crazy Egg). This makes it so that any system implementing Crazy Egg is limited in how large it can scale its operations (Crazy Egg). This software also requires that all data is stored and handled by Crazy Egg, making it difficult for Capital One to organize and manage data, as well as risks the security of the data by not being able to handle it on their own servers (Crazy Egg). This can also be very costly when exceeding the limit and losing a customer's personal information.

Lastly, Hotjar is another tool that, like the other softwares Crazy Egg and Glassbox, has heatmaps and a session recording (Contentsquare). Hotjar also allows for mouse tracking, which would allow for more data to be gleaned, which could assist in web design and workflow (Contentsquare). To save storage space, Crazy Egg and Hotjar delete recorded sessions after a period of time (Contentsquare). This may pose a significant issue because if a bug is ongoing and is eventually investigated the recording may be removed.

Our tool aims to solve the issue of space efficiency, by storing clickstream data rather than having recordings which take up a significant amount of space. Our tool also won't require third party management, allowing us to handle the data internally. This way we ensure that we can control when we delete the data or know how it's used exactly. This will also allow us to manage the data and not risk the third party compromising our data.

In conclusion, our iteration of this technology will be designed to capture a user's browser session while still maintaining the security of any personal or sensitive information captured during the session, ensuring that the data is stored using minimal storage space, and being able to store all user sessions on Empath. We'll store our data on a cloud storage system to allow for it to be highly scalable, and it won't require any external software installation to obtain a user's session from the browser, as it will all be stored via the website's own clickstream data. Our final

project will accurately portray a user's journey through a website, while adhering to the requirements Capital One has set out.

Section B. Engineering Design Requirements

B.1 Project Goals (i.e. Client Needs)

Capital One's agents that use Empath frequently encounter issues that hinder their ability to help customers. The debugging process to remove these bugs is extensive as it takes a significant amount of time and effort to determine the bug's source, as there's currently no way to easily reproduce these bugs. To help replicate issues and debug the service Capital One needs a service that will:

- Improve the process of tracking and analyzing user interactions in a digital environment, enabling more detailed insights into agents' behaviors.
- Create a streamlined replay service that allows businesses to easily visualize agents' journeys.
- Establish a more scalable and efficient system for storing/replaying the recorded data(cloud services).
- Reduce the time it takes to locate encountered issues in Empath.

B.2 Design Objectives

Our team will attempt to solve Capital One's problem by taking clickstream data from Empath and converting that into a playback for developers to debug the service with. Our team's design will:

- This design will provide a detailed replay feature, allowing users to visually track every click and interaction in the agents' journey.
- The design will replay an agent's journey after at least 1 minute.
- The design will be achievable given the resources: Heap, Amplitude, and Google Analytics, and a $1000 budget.
- The design will be realistic in terms of system scalability, allowing it to handle large amounts of data without significant performance degradation.
- The design will be completed in the timeframe of one academic school year.

B.3 Design Specifications and Constraints

Design Constraints

The design of our agent replay system must will meet its objectives by adhering to the following constraints:

- The design has to adhere to the $1000 budget for development costs and cloud services, including the potential costs for Heap, Amplitude, and Google Analytics.
- The design has to be implemented into a cloud service to allow for easier data management and scalability.
- The design has to track a user's clickstream data and store it for future access.
- The design must follow a project timeline that allows for full testing and completion within one academic year (9 months), with definable milestones.

Realistic Constraints

To ensure our product will be productive to those who use it, and to ensure its longevity, it will have the following realistic constraints:

- The design has to be seamlessly integrated with the platforms that the client is using without needing any significant infrastructure changes
- The design has to remain cost effective to match the allotted budget.
- The design must comply with data privacy regulations to ensure sensitive data is stored securely
- The design must be intuitive enough that non-technical personnel could operate it easily without any training or support.
- The design must reconstruct a journey only from collected clickstream data
- Ensure that it is scalable, with possibly many users accessing the database we want to make sure it still functions with a high load of users

B.4 Codes and Standards

    To ensure our project abides by our above requirements and rules of design, we'll abide by several standards and codes while developing our project. These standards and codes will be:

Standards
- ISO/IEC 27001: Information Security Management Systems (ISMS)
    - Our design will ensure that the system handling replay data maintains confidentiality and integrity of sensitive user data.
- IEEE 829-2008: Standard for Software and System Test Documentation
    - The design will undergo testing. This will allow for documentation of the test results, allowing for future improvements and maintenance.
- ISO 9241-210: Ergonomics of Human-System Interaction
    - The design will be intuitive, user-friendly, and accessible, reducing user confusion and improving the overall user experience without need for extensive training/support to operate.
- IEEE 7000-2021
    - This design should have consideration of individual and ethical values. Such as, privacy, fairness, and accountability of user's data.

Codes
- FTC Act: Section 5 - Unfair or Deceptive Acts or Practices
    - The design must be transparent in its data collection and usage to avoid deceptive practices.
- ECPA (Electronic Communications Privacy Act)
    - Since the design involves tracking communications between the agent and customer, it must comply with ECPA to ensure privacy in electronic communications.
- GLBA(Gramm-Leach-Bliley Act)
    - Requires financial institutions to explain their information sharing practices to their customers and to safeguard their data

Section C. Scope of Work

C.1 Deliverables

Sponsor Deliverables:
- Functioning Prototype of Agent Replay
- Playwright program capturing user clickstream data to "replay" their path through a website.
- Technical and User Documentation detailing how it works/how to use
- Fall poster

Academic Deliverables:
- Team Contract   (Sep. 6)
- Project Proposal   (Oct. 11)
- Fall Design Poster   (Nov. 15)
- Preliminary Report   (Dec. 9)
- Capstone Poster and Abstract (March 28)
- Final Design Report (May 2)

Obstacles:

- Google analytic clickstream may not be usable at first we may have to convert the software we use to get clickstream data.
- Playwright may not provide the right solution maybe it does live calls, which is what we don't want for security reasons
- Open source website doesn't allow for clickstream data to be collected
- Converting clickstream data to playwright may cause issues

C.2 Milestones

Milestone	Date/Time
Presentation Completion	Oct 3, 2024
Proposal Completion	October 11, 2024
Add Clickstream Events To Website	November 1, 2024
Store And Read Clickstream Data	November 15, 2024
Implementation of Click Stream Data Into Playwright	December 6, 2024
Functioning Prototype	Jan 13, 2025
Create Tests to Verify Software Behavior	Jan 27, 2025

Creation of User Documentation	Feb 10, 2025
Internal Testing/Review	March 3, 2025
Refinement	April 2, 2025
Capstone EXPO	April 26, 2025
Final Report	May 2, 2025

C.3 Resources

Our team's project's resources will be composed mostly of open-source or free software that we'll need to implement into our design, with a budget that we may use if the need arises. The resources required for this project would be:

Software

- IDEs for developing (VS Code)
- Playwright for Frontend Automation
- Version Control system(Github)
- Google Analytics API
- Cloud Services to store data and processing
- Mock data generation tools (Tweak)
- Relevant libraries for data handling
- Communication platform for collaboration
- Website Hosting

Budget

- $1000

Appendix 1: Project Timeline

Provide a Gantt chart of similarly composed visual timeline showing the start and end dates of all completed tasks and how they are grouped together, overlapped, and linked together. Include all senior design requirements including design reports and Expo materials (i.e. Abstract, Poster, and Presentation). All major milestones should be included in the timeline.
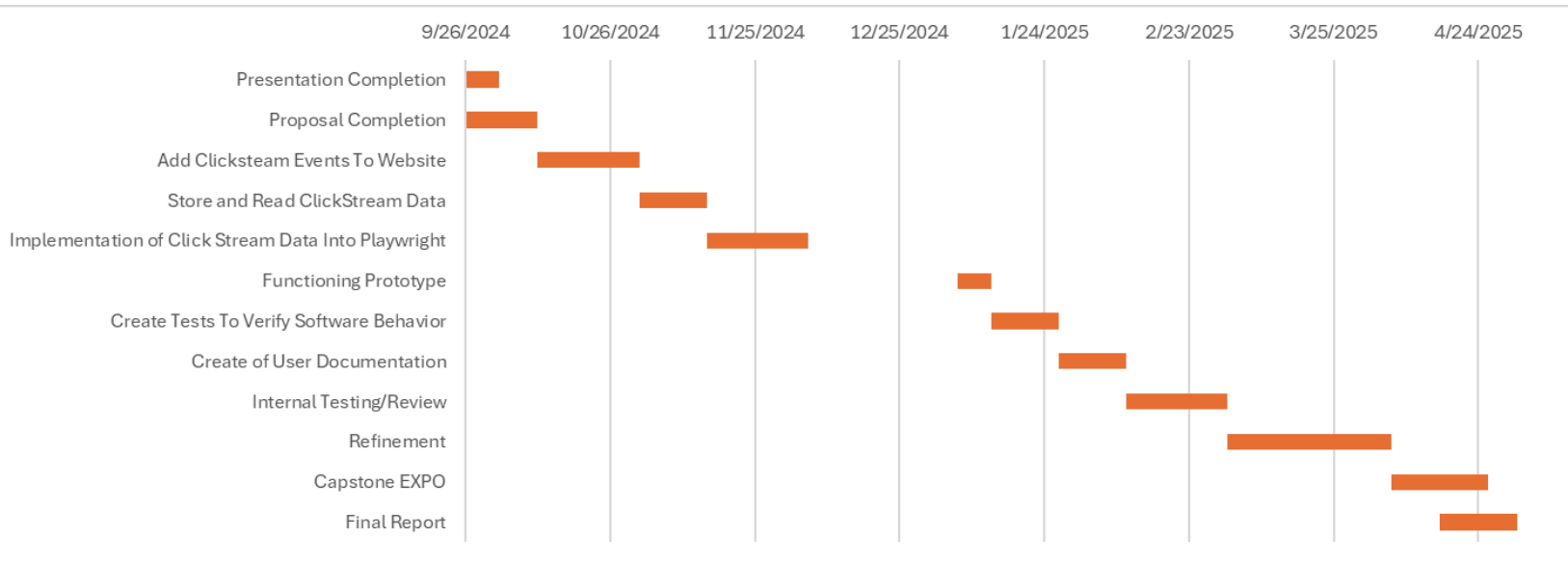
Appendix 2: Team Contract (i.e. Team Organization)

Step 1: Get to Know One Another. Gather Basic Information.

Task: This initial time together is important to form a strong team dynamic and get to know each other more as people outside of class time. Consider ways to develop positive working relationships with others, while remaining open and personal. Learn each other's strengths and discuss good/bad team experiences. This is also a good opportunity to start to better understand each other's communication and working styles.

Team Member Name	Strengths each member bring to the group	Other Info	Contact Info
Del Huband	Creative, Problem Solving, Strategic	I like being able to improve my skills as a programmer.	hubanddh@vcu.edu (804) 401-9219
Parker Gonzaga	Problem Solving, organization, some database experience	I enjoy working as a team and completing tasks efficiently. Willing to learn new techniques in order to fulfill project requirements.	gonzagapv@vcu.edu (571) 524-0964
Alex Nguyen	A bit of industry experience, problem solving, communication	I'm interested in AI.	nguyena57@vcu.edu (571) 340-0035
Brian Vo	Communication, analytical thinking	I like to develop things that have practical uses.	vobq2@vcu.edu (571) - 527 - 8826

Other Stakeholders	Notes	Contact Info
Rachita Sowle (Advisor)		sowler@vcu.edu
Mahesh Nair (Sponsor)		mahesh.bahulleyannair@capitalone.com
Derrick Murry (Sponsor)		derrick.murry@capitalone.com

Step 2: Team Culture. Clarify the Group's Purpose and Culture Goals.

Task: Discuss how each team member wants to be treated to encourage them to make valuable contributions to the group and how each team member would like to feel recognized for their efforts. Discuss how the team will foster an environment where each team member feels they are accountable for their actions and the way they contribute to the project. These are your Culture Goals (left column). How do the students demonstrate these culture goals? These are your Actions (middle column). Finally, how do students deviate from the team's culture goals? What are ways that other team members can notice when that culture goal is no longer being honored in team dynamics? These are your Warning Signs (right column).

Resources: More information and an example Team Culture can be found in the Biodesign Student Guide "Intentional Teamwork" page (webpage | PDF)

Culture Goals	Actions	Warning Signs
Being on time to every meeting	- Set up meetings in shared calendar - Ensure all parties have date cleared prior to setting meetings	- Student misses meetings regularly - Consistently missing meeting is brought up to advisor
Informing the group of any delays in completing assignments	- Stay up to date with each other's project responsibilities - Set reasonable deadlines and note when an extension is needed	- Student shows up for weekly meeting with no considerable work done
Accomplishing goals ahead of time	- Set deadlines - Work at a pace that would leave leeway before deadline	- Procrastinating - Missing deadlines
Delegating roles and understanding your job	- Assign each group member a task - Write out clear description for each task	- Unfinished work on any group member's end, causing project progress to slow down
Open Communication	- Don't be afraid to ask questions	- Student ceases communication with group

Step 3: Time Commitments, Meeting Structure, and Communication

Task: Discuss the anticipated time commitments for the group project. Consider the following questions (don't answer these questions in the box below):

- What are reasonable time commitments for everyone to invest in this project?
- What other activities and commitments do group members have in their lives?
- How will we communicate with each other?
- When will we meet as a team? Where will we meet? How Often?
- Who will run the meetings? Will there be an assigned team leader or scribe? Does that position rotate or will same person take on that role for the duration of the project?

Required: How often you will meet with your faculty advisor advisor, where you will meet, and how the meetings will be conducted. Who arranges these meetings?
See examples below.

Meeting Participants	Frequency Dates and Times / Locations	Meeting Goals Responsible Party
Students Only	As Needed, On Discord Voice Channel	Update group on recent challenges and accomplishments (Brian will scribe important details from the meeting)
Students Only	Every Thursday 5-7 in West Eng. 101	Actively work on project with team and communicate weekly updates and set future goals (Del will write out upcoming goals and other information from this meeting)
Students + Faculty advisor	Currently planned to meet weekly or bi-weekly with our advisor.	Update faculty advisor and get answers to our questions (Del will scribe the important details from this meeting)
Project Sponsor	Currently planned every Wednesday. (Can be increased/ adjusted)	Update project sponsor and make sure we are on the right track (Brian will scribe important details from the meeting)

Step 4: Determine Individual Roles and Responsibilities

Task: As part of the Capstone Team experience, each member will take on a leadership role, ***in addition to*** contributing to the overall weekly action items for the project. Some common leadership roles for Capstone projects are listed below. Other roles may be assigned with approval of your faculty advisor as deemed fit for the project. For the entirety of the project, you should communicate progress to your advisor specifically with regard to your role.

- **Before meeting with your team**, take some time to ask yourself: what is my "natural" role in this group (strengths)? How can I use this experience to help me grow and develop more?
- **As a group,** discuss the various tasks needed for the project and role preferences. Then assign roles in the table on the next page. Try to create a team dynamic that is fair and equitable, while promoting the strengths of each member.

Communication Leaders

Suggested: Assign a team member to be the primary contact <u>for the client/sponsor</u>. This person will schedule meetings, send updates, and ensure deliverables are met.

Suggested: Assign a team member to be the primary contact <u>for faculty advisor</u>. This person will schedule meetings, send updates, and ensure deliverables are met.

Common Leadership Roles for Capstone

1. **Project Manager:** Manages all tasks; develops overall schedule for project; writes agendas and runs meetings; reviews and monitors individual action items; creates an environment where team members are respected, take risks and feel safe expressing their ideas.
   Required: On Edusourced, under the Team tab, make sure that this student is assigned the Project Manager role. This is required so that Capstone program staff can easily identify a single contact person, especially for items like Purchasing and Receiving project supplies.
2. **Logistics Manager:** coordinates all internal and external interactions; lead in establishing contact within and outside of organization, following up on communication of commitments, obtaining information for the team; documents meeting minutes; manages facility and resource usage.
3. **Financial Manager:** researches/benchmarks technical purchases and acquisitions; conducts pricing analysis and budget justifications on proposed purchases; carries out team purchase requests; monitors team budget.
4. **Systems Engineer:** analyzes Client initial design specification and leads establishment of product specifications; monitors, coordinates and manages integration of sub-systems in the prototype; develops and recommends system architecture and manages product interfaces.
5. **Test Engineer:** oversees experimental design, test plan, procedures and data analysis; acquires data acquisition equipment and any necessary software; establishes test protocols and schedules; oversees statistical analysis of results; leads presentation of experimental finding and resulting recommendations.
6. **Manufacturing Engineer:** coordinates all fabrication required to meet final prototype requirements; oversees that all engineering drawings meet the requirements of machine shop or vendor; reviews designs to ensure design for manufacturing; determines realistic timing for fabrication and quality; develops schedule for all manufacturing.

Team Member	Role(s)	Responsibilities
Del Huband	Test Engineer	✔ Extensively test all code to ensure it operates as intended ✔ Analyzes code efficiency ✔ Ensures programs don't have significant overhead
Parker Gonzaga	Financial Manager	✔ Keep track of expenses ✔ Handle purchase requests ✔ Ensure expenses stay within project budget
Brian Vo	Project Manager	✔ Contact for sponsor ✔ Keep group on track, track agendas, schedule meetings
Alex Nguyen	Systems Engineer	✔ Analyze client's initial design specifications ✔ Establishes product specifications ✔ Develops system architecture

Step 5: Agree to the above team contract

Team Member: Del Huband   *Signature: Del Huband*

Team Member: Alex Nguyen   *Signature:* **Alex Nguyen**

Team Member: Parker Gonzaga   *Signature: Parker Gonzaga*

Team Member: Brian Vo   *Signature: Brian Vo*

References

[1] Contentsquare. (n.d.). *Hotjar*. Retrieved October 6, 2024. https://www.hotjar.com/

[2] Glassbox. (n.d.). *Glassbox*. Retrieved October 7, 2024. https://www.glassbox.com/

[3] Crazy Egg. (n.d.). *Crazy Egg*. Retrieved October 10, 2024. https://www.crazyegg.com/

[4] Bekmambetov, T. Bedford, G. Kehoe, D. 2019. Method for recording, editing and reproduction of computer session (European Patent 3457295A2). European Patent Office. https://patents.google.com/patent/EP3457295A2/en?q=(%22browser+playback%22)&oq=%22browser+playback%22