

# Reflection and Traceability Report on Plutos

Team #10, Plutos

Payton Chan

Eric Chen

Fondson Lu

Jason Tan

Angela Wang



# 1 Changes in Response to Feedback

## 1.1 SRS and Hazard Analysis

Table 1: SRS Issues

Issue Number	Source	Issue Title	Addressed?	PR	Comments
<a href="#">105</a>	TA	docs(SRS): Include formal/math specs	✓	<a href="#">300</a>	Added formal math specifications in SRS for OCR model, data constraints, budget calculation and receipt/receipt item data definitions
<a href="#">106</a>	TA	docs(SRS): Add section for "normal operation"	✓	<a href="#">317</a>	Added section for normal operation and use case diagram
<a href="#">107</a>	TA	docs(SRS): Add section for undesired event handling	✓	<a href="#">308</a>	Added section
<a href="#">108</a>	TA	docs(SRS): Add fit criteria to requirements	✓	<a href="#">310</a>	Added fit criteria
<a href="#">109</a>	TA	docs(SRS): Link requirements to rationale section	✓	<a href="#">303</a>	Linked requirements to rationale section
<a href="#">110</a>	TA	docs(SRS): Link likely and unlikely changes to requirements	✓	<a href="#">303</a>	Linked requirements to un/likely changes
<a href="#">38</a>	Peer Review	peer-review[team 23]: OCR system accuracy	✓	—	This was a question, which was answered as a comment in the issue.
<a href="#">39</a>	Peer Review	peer-review[team 23]: OCR pipeline privacy	✓	—	This was a question, which was answered as a comment in the issue.
<a href="#">40</a>	Peer Review	peer-review[team 23]: release platform	✓	<a href="#">310</a>	Changes made
<a href="#">41</a>	Peer Review	peer-review[team23]: Non-Functional Requirements Verifiability	✓	<a href="#">179</a>	Added details to requirements
<a href="#">42</a>	Peer Review	peer-review[team 23]: Ambiguous Response Time Specifications	✓	<a href="#">179</a>	Added details to specifications
<a href="#">43</a>	Peer Review	peer-review[team 23]: Item Recognition and Categorization Requirements	✓	<a href="#">299</a>	Slightly reworded requirements to make them more clear, but the issue was more of a clarification question than feedback
<a href="#">44</a>	Peer Review	peer-review[team 23]: Data retention and Deletion Policies	✓	<a href="#">179</a>	Added more details on policies

Table 2: Hazard Analysis Issues

Issue Number	Source	Issue Title	Addressed?	PR	Comments
<a href="#">111</a>	TA	docs(hazards): Add list of tables	✓	<a href="#">288</a>	Added table
<a href="#">112</a>	TA	docs(hazards): Put constants in constants section	✓	<a href="#">288</a>	Referred to symbolic constants
<a href="#">113</a>	TA	docs(hazards): Fix hazard recommended action	✓	<a href="#">299</a>	Changed recommended action
<a href="#">51</a>	Peer Review	Peer Review (hazards) - Expand Analysis on External System Interactions	✓	<a href="#">311</a>	Added assumption for external systems
<a href="#">52</a>	Peer Review	Peer Review (hazards) - Enhance Network Failure Handling and Data Integrity Measures	✓	<a href="#">185</a>	Added more details
<a href="#">53</a>	Peer Review	Peer Review (hazards) - Improve Hazard Mitigation User Feedback	✓	—	Was already addressed, left comment on issue
<a href="#">54</a>	Peer Review	Peer Review (hazards) - some assumptions needed for user end equipment	✓	<a href="#">187</a>	Added suggested assumptions
<a href="#">55</a>	Peer Review	Peer Review (hazards) - concern about problem resolution	✓	<a href="#">187</a>	Added suggested changes



## 1.2 Design and Design Documentation

Table 3: Design Doc Issues

Issue Number	Source	Issue Title	Addressed?	PR	Comments
<a href="#">233</a>	TA	docs(design): add tests to CI	✓	<a href="#">262</a> , <a href="#">264</a> , <a href="#">258</a>	Added tests to CI
<a href="#">234</a>	TA	docs(design): DetDesDoc syntax - change naming and data types	✓	<a href="#">247</a>	Added suggested changes
<a href="#">235</a>	TA	docs(design): DetDesDoc syntax - add transaction class	✓	<a href="#">313</a>	Updated design docs
<a href="#">236</a>	TA	docs(design): DetDesDoc semantics - add more details to ML/categorization module	✓	<a href="#">247</a>	Added details
<a href="#">237</a>	TA	docs(design): DetDesDoc semantics - add details about calculations	✓	<a href="#">247</a>	Added math specification
<a href="#">238</a>	TA	DetDesDoc semantics - add state diagram	✓	<a href="#">302</a>	Added state diagram to Figma
<a href="#">239</a>	TA	DetDesDoc semantics - add 'main' UI module	✓	<a href="#">302</a>	Adjusted DAG diagram and added Main UI module to module hierarchy and DAG
<a href="#">240</a>	TA	DetDesDoc semantics - exception handling specs	✓	<a href="#">302</a>	Added exception handling section for item misclassification
<a href="#">158</a>	Peer Review	Module hierarchy diagram Link Issue	✓	<a href="#">302</a>	Adjusted DAG diagram
<a href="#">159</a>	Peer Review	OCR Processing Module Issue	✓	<a href="#">302</a>	Added exception handling section for OCR module
<a href="#">160</a>	Peer Review	Peer Review - Limited Discussion on Module-Level Secrets	✓	<a href="#">313</a>	Added module secrets to image processing and categorization modules
<a href="#">161</a>	Peer Review	Peer Review - Limited Error Handling Descriptions	✓	<a href="#">302</a>	Added more details on exception handling
<a href="#">162</a>	Peer Review	Peer Review - Ambiguity in Access Routine Semantics	✓	<a href="#">320</a>	Added some more details on state. Many of the modules do not have an attached state though so they do not have the pre/post conditions set.
<a href="#">163</a>	Peer Review	Peer Review - Hyperlink in section 5	✓	<a href="#">186</a>	Fixed



### 1.3 VnV Plan and Report

Table 4: VnV Plan Issues

Issue Number	Source	Issue Title	Addressed?	PR	Comments
<a href="#">118</a>	TA	docs(vnv): add captions for tables	✓	<a href="#">241</a>	Added captions
<a href="#">119</a>	TA	docs(vnv): add to testing plan	✓	<a href="#">290</a>	Adjust functional testing criteria.
<a href="#">120</a>	TA	docs(vnv): system tests for FRs	✓	<a href="#">290</a>	Adjust control of system tests.
<a href="#">121</a>	TA	docs(vnvplan): specify usability survey as extra	✓	—	Switched Extra2 to user manual
<a href="#">71</a>	Peer Review	Peer Review - Acronyms table missing content	✓	<a href="#">188</a>	Added missing acronyms
<a href="#">72</a>	Peer Review	Peer Review - Enhance Usability and OCR Accuracy Test Scenarios	✓	<a href="#">188</a>	Added details
<a href="#">73</a>	Peer Review	Peer Review - Incomplete Error Handling and Recovery Test Cases	✓	<a href="#">194</a> , <a href="#">197</a>	Added more details
<a href="#">74</a>	Peer Review	Automated Testing and Verification Tools	×	—	Daily sanity checks aren't necessary as we are running the tests and linter pipelines on every PR. This will ensure that the app/features work seamlessly after every change.
<a href="#">75</a>	Peer Review	Load Testing for Concurrent Users	✓	—	Already addressed in initial version of VnVPlan 'Performance tests can be conducted to measure the app's speed and reliability, especially when processing large receipts or handling multiple users'.
<a href="#">85</a>	Peer Review	Peer Review - Enhance Testing for Data Persistence During App Crashes	✓	<a href="#">312</a>	Added Usability Test to cover data retention during unexpected crashes.



Table 5: VnV Report Issues

Issue Number	Source	Issue Title	Addressed?	PR	Comments
<a href="#">304</a>	TA	docs(vnv report): add more details to verifiability of tests	✓	<a href="#">316</a>	Add specific test details
<a href="#">265</a>	Peer Review	Peer Review - The lost xlsx file in hyperlink	✓	<a href="#">291</a>	Fixed
<a href="#">266</a>	Peer Review	Missing Functionality	✓	<a href="#">306</a>	A couple of the FRs weren't implemented yet during the initial draft of VnVReport - they have now been implemented. As for account update and notifications, we felt that for the scope of the project, these functionalities would be able to be implemented later.
<a href="#">267</a>	Peer Review	Reduced accuracy	✓	<a href="#">306</a>	Added more reasoning as to why we decreased accuracy threshold.
<a href="#">268</a>	Peer Review	Reduced accuracy	✓	<a href="#">316</a>	Multiple NFRs have already been tested (NFR-ACC-1, NFR-ACC-3, NFR-MTB-1) before this issue. Added details to some more NFRs in new PR.
<a href="#">269</a>	Peer Review	Lack of Detailed Analysis for Performance Test Omissions	✓	<a href="#">318</a>	Added rationale to test case omission
<a href="#">270</a>	Peer Review	Unclear Implications of Security and Legal Compliance Testing	✓	<a href="#">309</a>	Added more details

## 2 Challenge Level and Extras

### 2.1 Challenge Level

The expected challenge level is **general**. The primary challenge of the project is developing a machine learning model that can accurately parse items from a picture of a receipt, and to categorize them into appropriate spending categories. This requires a strong understanding of training and tuning models on image data to achieve high accuracy. Additionally, different items across various stores may have similar names or be difficult to recognize, which adds to the complexity of the task. The other part of the project is to develop a user-friendly mobile application, which is a more general software engineering component.

### 2.2 Extras

1. **Requirements elicitation report:** We have conducted interviews and a survey to gather requirements from potential users to determine user needs and preferences, and document the findings in a report. See [Requirements Elicitation Report](#).
2. **User manual:** We have created a user manual to help users understand how to use the application and its features. See [User Manual](#).

## 3 Design Iteration (LO11 (PrototypeIterate))

From the get go, Plutos was designed around allowing students to scan receipts using OCR to automatically log their expenses. Our early prototypes validated this concept, but through iterative testing and feedback, we realized that effective receipt scanning was just one part of the budgeting experience. The surrounding user interface and workflow had to evolve to support that feature in a way that felt intuitive, helpful, and student-friendly.

Initially, the app's layout was minimal — receipts were scanned and stored in a chronological list, with categorized expenses displayed. However, we found that while the scanning feature worked well, users often felt disconnected from their overall budget. They wanted to immediately understand how a new purchase impacted their spending limits. In response, we redesigned the post-scan experience to include a quick budget summary and an intuitive workflow to allow them to see their spending trends, budget accordingly and also view all previous expenses regarding certain spending categories.

Visually, we evolved from a generic, grey-toned interface to a more colorful, student-friendly aesthetic. This wasn't just for looks — feedback suggested that students responded better to a more engaging interface, especially one with positive reinforcement (e.g. green progress bars, colourful graphs, etc.).

A few notable changes we made were decreasing the OCR accuracy constraint, and leaving out push notifications in the initial version of the app. The OCR accuracy constraint was decreased to allow for a wider range of receipts

to be scanned and after testing, we found that the parsed receipt data was still accurate enough to be useful to users. We also decided to leave out push notifications in the initial version of the app because we already implemented in-app alerts once users input expenses that would exceed 80% of their budget. This was a conscious decision to avoid overwhelming users with notifications, and we didn't want to add stress to users who are already managing their finances. We communicated this decision with users and through feedback, they expressed that even though they were slightly disappointed, they understood the reasoning behind our design decision and also expressed that not having push notifications wasn't a deal breaker for them.

Throughout the design process, we gathered feedback from potential users and the team to iteratively improve the design of the app. With each iteration, decisions were made with our target users in mind to identify pain points early regarding budget visibility, personalization and overall design.

## 4 Design Decisions (LO12)

### Assumptions

Early in the design process, we assumed students would be highly motivated to categorize every expense and analyze their spending in detail. However, user feedback showed that most students wanted to engage with the app quickly and only when necessary. This challenged us to shift our focus away from detailed manual categorization and toward quick, automatic summaries and passive insights.

### Constraints

The OCR accuracy constraint was decreased to enhance the model's generalization, which in turn reduces ambiguity during data classification. A lower accuracy threshold encourages a more flexible model that avoids overfitting to the training data and can better adapt to unseen, real-world inputs. This is especially important when dealing with noisy, inconsistent, or highly varied data such as scanned receipts or user-generated content, where exact matches are less common. For the scope of this project, the team decided that by focussing on balancing accuracy and generalization, the model can remain lightweight and responsive while still offering meaningful performance. Additionally, the ability for users to modify misclassified entries will further help address any discrepancies and improve the overall accuracy over time through user feedback and potential retraining mechanisms.

### Limitations

We decided to leave out push notifications in the initial version of the app since we alert users once they've inputted an expense that exceeds 80% of their

budget in-app. This decision was made because we wanted to avoid overwhelming users with notifications, especially since they are already receiving in-app alerts.

Striving for an extremely high accuracy and implementing push notifications could lead to longer development times with disproportionate benefits, so we decided to prioritize a more streamlined and user-friendly experience over these features for the initial implementation.

## 5 Economic Considerations (LO23)

There is a clear growing market for budgeting and personal finance tools, especially amongst university students and young adults. Young adults often face financial challenges during their early years of living independently, and are also often unaware of the costs of living, driving the need for budgeting tools. However, being university students and young adults ourselves, we understand the barrier to entry to start budgeting is high. As well, current budgeting tools are often too complex and overwhelming for young adults to use, which is what drives the need for Plutos – a simple yet effective way to track expenses and budget using receipt photo scanning.

Marketing Plutos would involve a combination of digital outreach and grassroots efforts. This could include social media advertising (especially on platforms like Instagram and TikTok), collaborations with student organizations, and targeted campaigns during back-to-school periods. Additionally, in-person promotion at campus events and integration with student services at universities could help increase visibility and adoption.

Since the overhead cost of developing and maintaining Plutos is low, we would host the product on Google Play Store and Apple App Store for free because we understand students wouldn't want to pay for a budgeting tool. However, we would offer a premium tier to allow users to utilize additional features such as advanced analytics, bank account integration, budgeting suggestions and exporting budgeting data to a PDF or CSV. We estimate the cost of premium services to be \$1.99/month and with 105 monthly users, we'd be able to make money and cover the costs of maintaining the app (since the app is low cost and will only require \$210/month to operate).

We believe this target is feasible as we have already surveyed 20 students and all of them expressed interest in using Plutos. With a potential user base of 40,000 students just at McMaster University alone and our marketing plan, we believe that everyone can use Plutos immediately.

## **6 Reflection on Project Management (LO24)**

### **6.1 How Does Your Project Management Compare to Your Development Plan**

Our team followed the development plan as outlined in the SRS Development Plan section. We adhered to the planned meetings, communication methods, assigned roles, and workflow structure. The project was structured into phases, and we prioritized high-priority functional requirements as scheduled. While some deadlines were subject to change, we adapted accordingly to ensure steady progress.

At the start of the project, we had a general idea of using an OCR model to extract data from receipts, but we weren't committed to a specific tool. Throughout the development process, we experimented with a few different OCR models to see which one would work best for our needs. After testing and comparing results, we ultimately decided to use Tesseract, which is an open-source OCR engine developed by Google.

Tesseract provided the balance we were looking for in terms of accuracy and ease of integration. While some of the other models we tried had certain strengths, Tesseract performed consistently across a wide range of receipt types, and it was easier to work with in our app's pipeline. Even though we didn't plan on using it from the beginning, the decision to pivot toward Tesseract ended up being a good one for the project.

### **6.2 What Went Well?**

Our capstone group worked well together, maintaining strong organization and collaboration throughout the project. Meetings and clear communication ensured that everyone stayed on track with their tasks. Each team member effectively contributed to their assigned roles, which helped keep development on schedule.

A key success was the OCR model used to parse receipts. It met our expected accuracy, allowing us to extract and categorize receipt data effectively. This was essential for the project's goal of creating an innovative expense-tracking budgeting app.

### **6.3 What Went Wrong?**

Like any project, we encountered some challenges along the way. There were occasional difficulties in coordinating schedules, especially when balancing coursework with project deadlines.

In terms of workflow, there were moments where certain tasks took longer than expected due to debugging and integration challenges. However, the team adapted and adjusted the timeline as needed.

In terms of tools, one of the main challenges was our initial unfamiliarity with OCR models. Since most of us had limited prior experience working with

OCR technology, there was a learning curve when it came to understanding how different models worked, how to train or fine-tune them, and how to handle various types of receipt formats. It took time to experiment with different models and figure out how to integrate them into our pipeline effectively.

## **6.4 What Would you Do Differently Next Time?**

For future projects, we would aim to refine our approach to time management by setting aside more buffer time for debugging and testing. While our communication was strong, we could improve further by implementing more structured progress tracking to ensure that any potential delays are addressed earlier.

# **7 Reflection on Capstone**

## **7.1 Which Courses Were Relevant**

Much of the knowledge that contributed to the development and completion of this project stemmed from the courses we took during our undergraduate studies. Several of these courses provided essential foundational knowledge, particularly our project courses such as ENGINEER 1P13 and 2PX3, which helped us establish a solid and efficient framework throughout the development process. Other courses supported us in the technical aspects of the project. All team members gained a fundamental understanding of machine learning through a recent elective, COMPSCI 4ML3, and we also applied insights from our software design and testing courses, which were SFWRENG 3A04 and 3S03 as well as our interface from SFWRENG 4HC3 to effectively build our system. Overall, various elements of our project were informed by the knowledge we acquired in other courses.

## **7.2 Knowledge/Skills Outside of Courses**

This project presented a new challenge for our entire team. Consequently, we approached the capstone as a valuable learning experience, recognizing our roles as both team contributors and students. Naturally, not all group members were fully acquainted with the technologies and systems involved in this project. For many, it was their first experience working with and training a real machine learning model. Additionally, developing a mobile application posed its own challenges, but our team demonstrated significant effort to overcome them. Throughout this project, we also cultivated important soft skills, such as effective collaboration and task delegation. The capstone represented a significant milestone in each of our professional careers and provided an excellent opportunity for learning.