

# CSC490 Assignment A1 - Project Landscape

Chris Jiang (Student ID: 1008109574) | Yihe Li (Student ID: 1009802331)  
Qianwen Wang (Student ID: 1009985358) | Seonghyun Ban (Student ID: 1005127738)

January 24, 2026

## 1 Interest Statements

Accounting is a critical operational function for any business, yet it is practically a foreign language to most founders and early-stage operators that are non-accountants. When starting a small business, the cognitive overhead of understanding double-entry bookkeeping, account types, debit/credit rules, accruals, depreciation, and corporate-specific mechanics (e.g., dividends, capital cost allowance, prorating expenses) often becomes a significant barrier to progress.

For many founders, accounting isn't just confusing, it's a constant mental tax that disrupts building the business itself. Founders want to focus on customers, product, and revenue, not categorizing transactions, remembering account types, or worrying if they recorded something wrong (which has real legal consequences if that happens). Our project aims to remove accounting from the founder's brain as much as possible by acting like an "outsourced accountant in software form": automatically posting correct journal entries in the correct accounts based on what the transactions are, reconciling balances, and to automate the process as much as possible (e.g. automatic recurring entries for prepaid and prorated expenses, and automatic recurring entries for asset depreciation).

### 1.1 Individual Contributions

- **Chris** wants to build the auto LLM-based bookkeeping pipeline with strong accuracy + confidence gating.
- **Yihe** wants to build receipt/invoice extraction + vendor intelligence to improve automation quality (better recognition of recurring transactions).
- **Qianwen** wants to build the review + reconciliation workflow so users barely touch accounting unless necessary (i.e. unless notified by the accounting AI that something needs explanation).
- **Seonghyun** wants to build the ledger correctness + audit trail system (undo, revisions, explainable logs, exports).

## 2 Landscape Analysis

Relevant Item	Description	Commentary
QuickBooks & Intuit Assist	Popular SMB accounting platform; Intuit Assist is positioned as a GenAI financial assistant embedded into QuickBooks.	Strong incumbent + existing “AI” layer. Gap: it’s mainly an <b>insight</b> tool, not an <b>action</b> tool.
Xero & JAX (“Just Ask Xero”)	Xero’s AI tooling includes an assistant direction (JAX) for embedded Q&A and guidance in-product.	Great reference for in-product assistant UX. Gap: doesn’t address the pain-points that we’re solving, namely how to correctly record and handle complex accounting situations and entries.
Pilot	Outsourced bookkeeping/tax/advisory for startups; combines software + humans.	Strong proof that founders want accounting “handled”, which demonstrates market demand. Gap: ultra-expensive and human-dependent; we aim for software-first and real-time guidance for smaller businesses and solopreneurs, at a very affordable price.
Vic.ai (accounts payable Automation)	AI-first accounts payable automation from invoice ingestion to payment execution.	Shows the enterprise-grade automation path. Gap: AP automation is NOT full accounting; we focus on general ledger correctness, beyond just payables.
Odoo Accounting	ERP accounting module supporting double-entry bookkeeping and accrual/cash basis functionality.	Great example of full-stack accounting workflows. Gap: powerful but still complex; requires extensive accounting knowledge to use.
GnuCash	Free and open-source personal/small business accounting software; supports double-entry and many reporting features.	Great baseline for “full-featured accounting” without AI. Gap: manual workflows + steep learning curve + extensive knowledge needed; we want “GnuCash correctness” with AI auto-drafting and guidance.

Relevant Item	Description	Commentary
Ledger CLI	Free and open-source command-line double-entry accounting using a text journal; entries must balance.	Excellent for verifiable bookkeeping and constraints. Opportunity: leverage this idea to build a balancing/validation core so the AI can't output nonsense entries.
[Research] Audit-Copilot (LLMs + fraud detection in bookkeeping)	Research exploring LLMs as anomaly detectors for double-entry bookkeeping/journal-entry testing.	Useful because it treats bookkeeping as a structured ledger reasoning task. Shows opportunity for LLMs beyond chat: detecting anomalies + sanity checking.
[Research] Automated Invoice Data Extraction (LLM + OCR)	End-to-end hybrid approach combining OCR, deep learning, and LLMs for invoice extraction.	Directly relevant to ingesting receipts/invoices. Gap: extraction is still not the same as posting entries; we connect extraction into the pipeline of: account mapping + posting journal entries + explainability.

### 3 Project Outline

#### 3.1 Problem Statement

Small business accounting becomes painful because founders must repeatedly make decisions that are too confusing to reason about, and take away time and mental energy from running their business, like:

- what account does this transaction concern?
- what are the rules on depreciating this asset?
- what's the correct debit/credit split? (especially in complex transactions)
- did I reconcile this correctly?
- what taxable events are triggered when I pay myself? what should I do now?

Even when software exists, it assumes that the user has accounting knowledge, and the founder remains the bottleneck.

#### 3.2 Proposed Solution

Build an AI Accounting Autopilot that performs bookkeeping end-to-end:

- Ingest transactions + receipts
- Generate journal entries automatically
- Reconcile balances
- Produce clean books continuously
- Ask for clarification only when needed

Goal: the founder spends as little time as possible on bookkeeping except rare exceptions.

#### 3.3 High Level System Design

Inputs (not exhaustive):

- Bank feed / uploaded statements
- Receipt & invoice uploads
- Business profile settings (industry, tax rules, etc.)

Pipeline:

1. **Ingestion + normalization**
  - Clean transaction text fields
  - Parse amounts, vendors, and recurring patterns
2. **Document understanding**
  - Receipt/invoice extraction: vendor, date, taxes, total, line items
3. **Posting engine (core ML task)**
  - Transaction → predicted accounts + journal entry structure
  - Supports split postings (e.g., tax + expense + fees)
4. **Hard validation rules (must-pass)**
  - Every entry balances (debits = credits)
  - No illegal account combinations
  - Required fields are present

## 5. Reconciliation layer

- Match postings to bank statements
- Flag unmatched transactions / duplicates

## 6. Confidence gating

- High-confidence → auto-post
- Low-confidence → ask *one* minimal clarification question

## 7. Audit trail + export

- Every entry has provenance (why/how it was generated)
- Export formats for accountants / taxes

### 3.4 Milestones

- M1: Define scope + dataset schema + base Chart of Accounts templates
- M2: Build ingestion pipeline (bank CSV + simple receipt OCR)
- M3: Baseline classifier (transaction → account category)
- M4: Journal entry builder + balancing validator
- M5: Explanation UI (“coach mode”) + feedback loop
- M6: Improve accuracy (active learning from user corrections)
- M7: End-to-end demo: “upload → draft books → explain → export”

### 3.5 Unknowns

- Correctness expectation is extremely high (finance + taxes)
- Edge cases: mixed personal/business, partial refunds, CCA-type asset purchases (although this can be avoided for MVP)
- Regional differences in business and tax law (Canada vs US)
- User trust: automation must be safe + reversible if requested

## 4 Press Release

### Introducing Autobook — The Accounting Autopilot for Small Businesses

**Toronto, ON — (January 24, 2026)** — Today we're announcing **Autobook**, an AI bookkeeping assistant for small business owners that removes the stress of accounting by turning messy bank transactions into clean bookkeeping records with minimal effort.

Most accounting tools still make founders do the hardest part: deciding *what a transaction is* (expense? asset? owner draw? tax?), then categorizing it correctly. Autobook changes that workflow: it **detects what each transaction is likely to represent**, asks a short clarification question only when uncertain, and then posts a correct bookkeeping entry automatically.

Autobook is designed for founders who don't want to learn accounting — they want to **run their business** without worrying that their books are wrong.

#### The Problem

Accounting is often a “silent tax” on early founders. Even when a business is doing well, bookkeeping creates constant friction:

- Transactions are unclear (e.g., “Amazon” could be supplies, equipment, or personal).
- Mistakes compound over time and make taxes stressful.
- Founders delay bookkeeping, leading to messy records and uncertainty.

#### The Autobook Solution

Autobook reduces bookkeeping to the smallest possible interaction:

##### 1. Detect

- Autobook reads bank transaction descriptions, amounts, dates, and receipt metadata.
- It predicts the top likely categories (e.g., subscription, supplies, revenue, asset purchase).

##### 2. Ask (only if needed)

- If the model is uncertain, Autobook asks a single, minimal question like:
  - “Was this *supplies* or *equipment*? ”
  - “Is this *business* or *personal*? ”
- The user answers in one click, without touching debits/credits or account names.

##### 3. Post & Remember

- Autobook records the appropriate bookkeeping entry using verified templates.
- It learns from the user's choice so the next similar transaction requires no question.

#### Key Features

- **Low-effort bookkeeping:** Autobook handles classification and posting with minimal user input.
- **Confidence-based questions:** It only interrupts when ambiguity is high.

- **Consistency over time:** Autobook adapts to the business (vendors, patterns, recurring bills).
- **Clean exports and audit trail:** Entries include traceable reasoning and can be exported for accountants.
- **Undo and corrections:** Users can fix mistakes, and the system updates future behavior automatically.

## What This Feels Like in Real Life

- **Stripe – \$49.00** → Autobook recognizes recurring software billing and posts it automatically.
- **Amazon – \$612.38** → Autobook asks: “*Supplies or Equipment?*” The user clicks once.
- **Next month Amazon – \$590.00** → Autobook remembers and handles it with no question.

## Fictional Quotes (Users & Stakeholders)

*“This is the first bookkeeping tool that doesn’t make me feel like I’m doing homework.*

*I just answer one question and move on.”*

— Small business founder

*“Autobook reduces the number of bookkeeping decisions a human needs to make.*

*That’s exactly how you scale clean financial records.”*

— CPA / Accounting advisor (stakeholder)

*“My books went from ‘I’ll do it later’ to ‘done automatically.’ It’s like having a bookkeeper who only pings me for edge cases.”*

— Online store owner

## Availability

Autobook will be released as a web app MVP supporting:

- Bank CSV import (with plans for direct bank integrations later)
- Receipt and invoice upload
- Transaction detection + clarification questions
- Bookkeeping entry export (for accountants and tax workflows)

## Appendix: Iterations Considered

### 1. Full Accounting Autopilot

- **Idea:** Autopost everything end-to-end with no user interaction.
- **Why we did not choose it:** Too risky for MVP trust; errors could be costly.

### 2. Accounting Tutor / Education Tool

- **Idea:** Teach users double-entry accounting and build understanding over time.

- **Why we did not choose it:** Many founders don't want to learn bookkeeping — they want it off their plate.

### 3. Chosen Direction: Detect → Ask → Post

- **Idea:** Predict what transactions represent, ask one question when uncertain, and post correctly.
- **Why it wins:** Minimal user effort, scalable accuracy via feedback, and safer than full autopilot.

*New platform enables building, discovering, sharing, and monetizing through constructing task structures—removing redundant problem-solving into compounding community knowledge.*