A Peek into Botkeeper's Machine Learning

AI UNCHAINED

Accounting Intelligence. No boundaries. No limits.

OUR TEAM

Brian Emposimato
Director of Product

NIshanth Vimalesh
Machine Learning Engineer

Jenny Blackwell
Director of Engineering

TABLE OF CONTENTS

AI Starts with a Business Problem

- ⬡   At Botkeeper, our goal is to automate and streamline as many steps in the accounting cycle as possible.
- ⬡   By dissecting these individual steps, we gain an understanding of the underlying challenges we need to solve.
- ⬡   As our understanding of a business problem gets deeper and more specific, we get a better understanding of how to apply AI to solve it.

...and Data

- Chart of Accounts
- Vendors and Customers
- Classes
- Departments
- Transactions
- etc.

Data-Layer
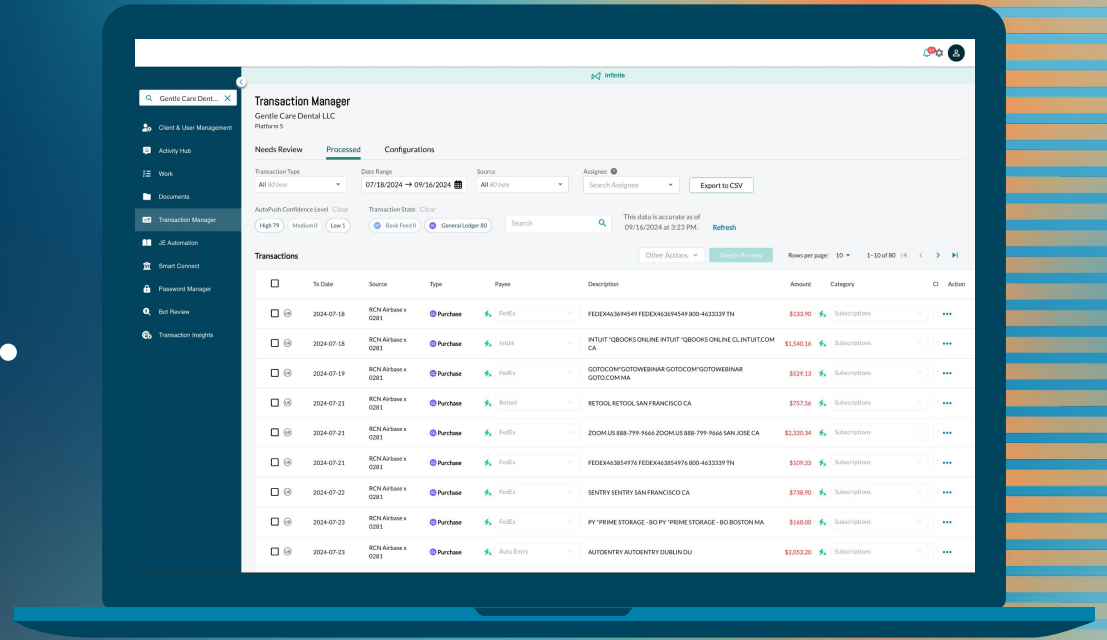
- Bank Feed
- Bank Statements

- Bank Statements

Have you ever thought you're spending too much time in the bank feed and there has to be a better way?

A Glance at
Transaction Manager

Transaction Manager combines machine learning with some light human assistance to auto-categorize transactions based on past transactions.

CATEGORIZATION

Leverage technology to categorize the transactions first and determine when human review is needed
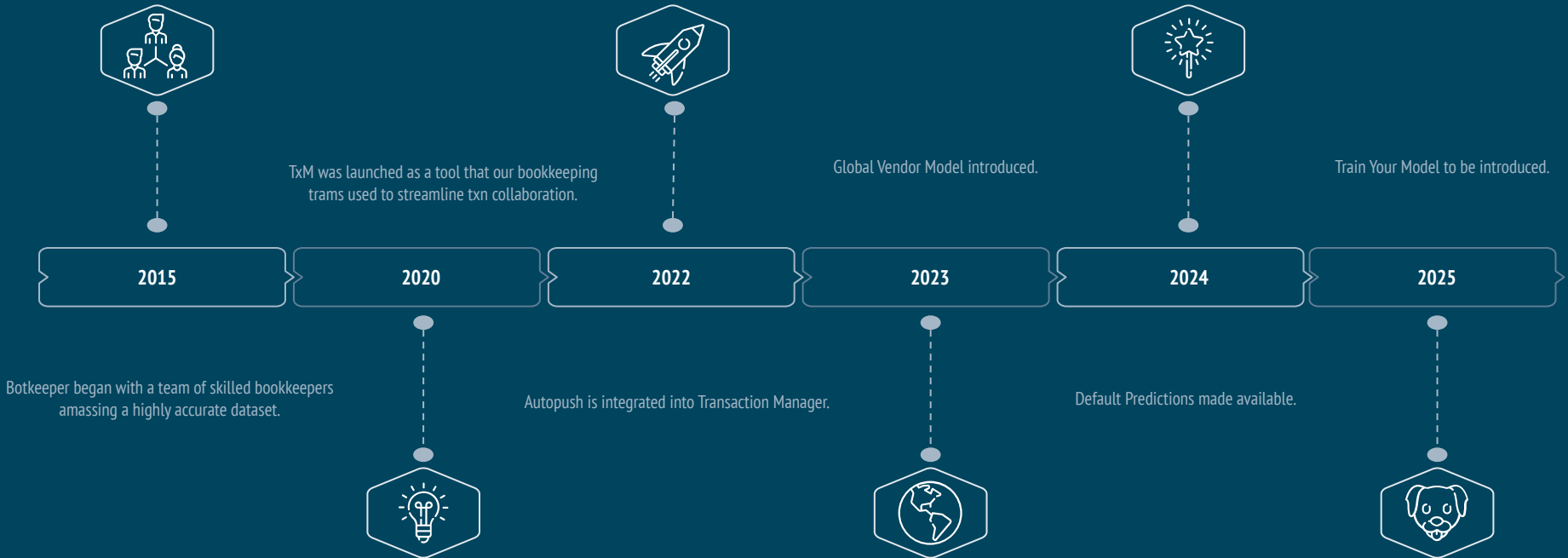
CUSTOMIZATION

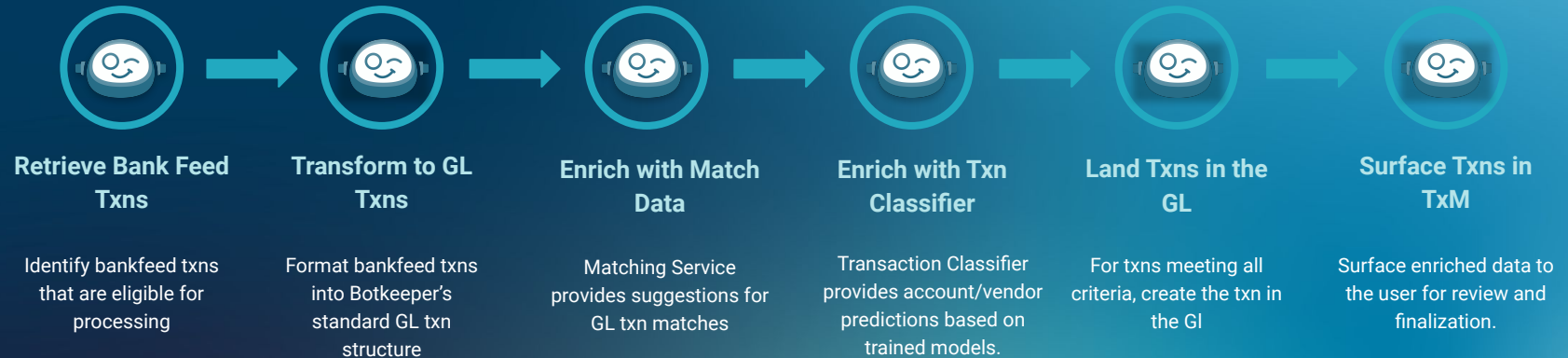Tailoring the experience to your client's specific needs

COLLABORATION

Streamline collaboration around collecting information on uncategorized transactions

Opportunities

The Evolution of Transaction Manager

TxM was launched as a tool that our bookkeeping trams used to streamline txn collaboration.

Global Vendor Model introduced.

Train Your Model to be introduced.

| 2015 | 2020 | 2022 | 2023 | 2024 | 2025 |

Botkeeper began with a team of skilled bookkeepers amassing a highly accurate dataset.

Autopush is integrated into Transaction Manager.

Default Predictions made available.

Transaction Manager and Autopush

Retrieve Bank Feed Txns

Identify bankfeed txns that are eligible for processing

Transform to GL Txns

Format bankfeed txns into Botkeeper's standard GL txn structure

Enrich with Match Data

Matching Service provides suggestions for GL txn matches

Enrich with Txn Classifier

Transaction Classifier provides account/vendor predictions based on trained models.

Land Txns in the GL

For txns meeting all criteria, create the txn in the GI

Surface Txns in TxM

Surface enriched data to the user for review and finalization.
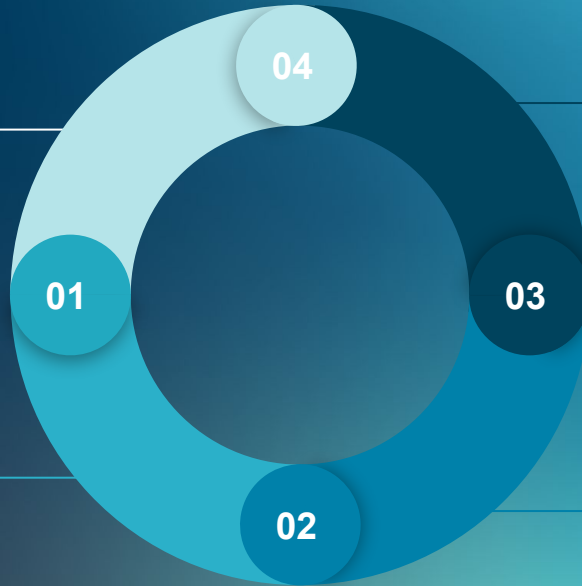
From Training Data to Trained Model

Collect Training Data

Our models learn from historically categorized transactions.

Model Training

Learning how important each attribute in our dataset is to solving our business problem.

Pre-processing and Enrichment

Pre-processing steps allow us to get the most useful information out of our data.

Feature Engineering

Feature engineering is a technique that leverages data to create new variables that aren't in the training data set.

04

01

03

02

EMBEDDING

- ⬡ Model does not understand words the way we do, they need numbers to work with.
- ⬡ Embedding is a way to turn words into a bunch of numbers, but in a smart way that captures their meaning and relationships.

amazon

amzn

walmart

amzn

TRANSFER LEARNING

- Each entity model starts off smarter, using knowledge already gained from the global model's embedding.
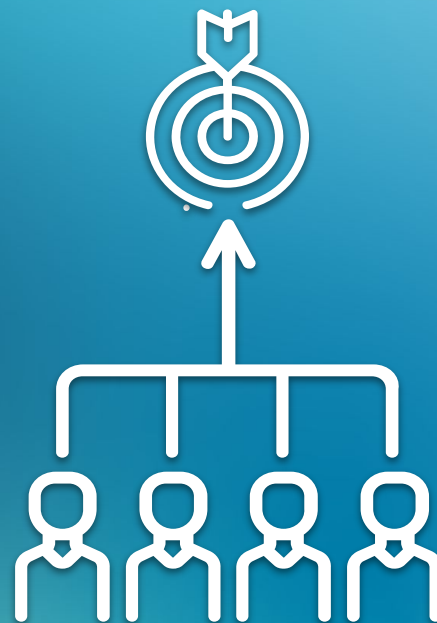- Allows our model to quickly adapt with fewer transactions and deliver more accurate results.

ENSEMBLE APPROACH

⬡ An ensemble model combines different models, each with their own strengths, to make better predictions than any one model could alone.

ENSEMBLE MODEL

SIMILARITY MODEL

Learns from past transactions

CLASSIFICATION MODEL

Learns from patterns in the transactions

GLOBAL MODEL

Learns from past global transactions

AI UNCHAINED

Accounting Intelligence. No boundaries. No limits.

AI & ML improve efficiency but never fully eliminate the need for humans. AutoPush can eliminate a majority of transaction categorization work and streamlines the rest.
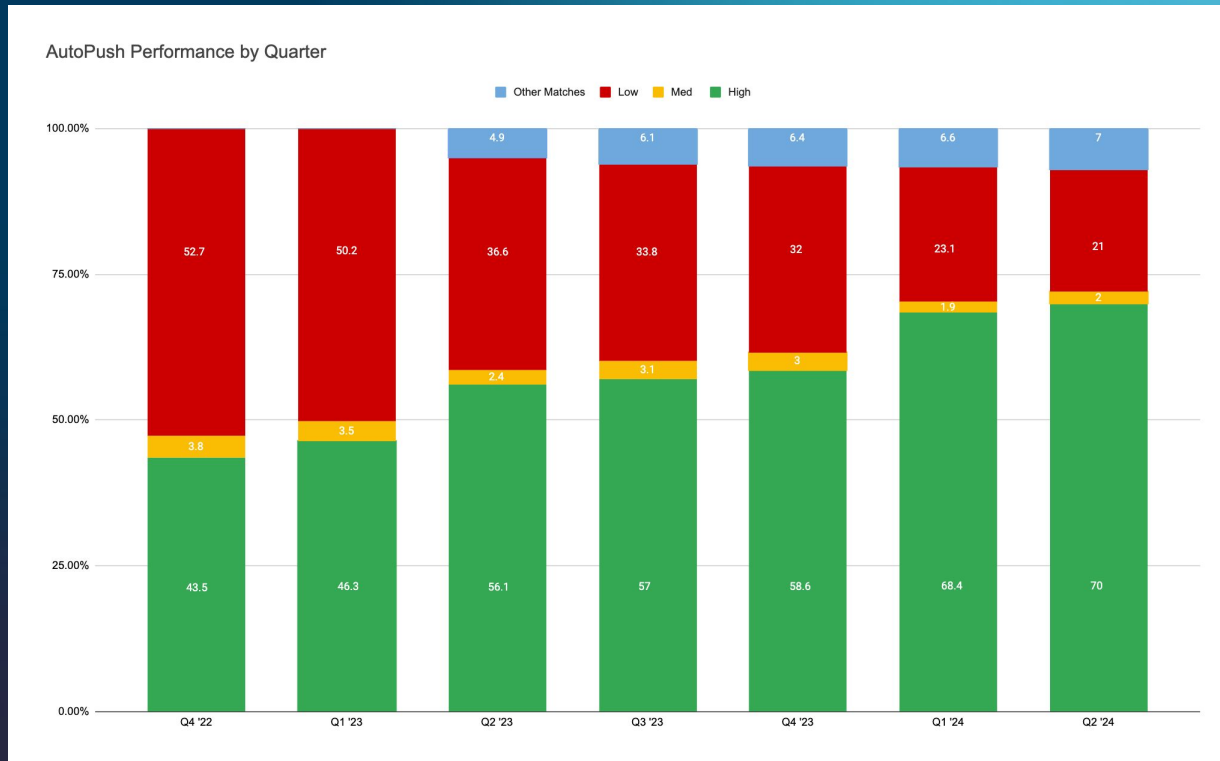
71%

Average Automated Transactions

AutoPush Metrics

AutoPush Transactions (V2) — Breakdown by Confidence Threshold - Total Transactions

AutoPush Metrics

AutoPush Performance by Quarter

Legend: Other Matches, Low, Med, High

Quarter	High	Med	Low	Other Matches
Q4 '22	43.5	3.8	52.7	
Q1 '23	46.3	3.5	50.2	
Q2 '23	56.1	2.4	36.6	4.9
Q3 '23	57	3.1	33.8	6.1
Q4 '23	58.6	3	32	6.4
Q1 '24	68.4	1.9	23.1	6.6
Q2 '24	70	2	21	7
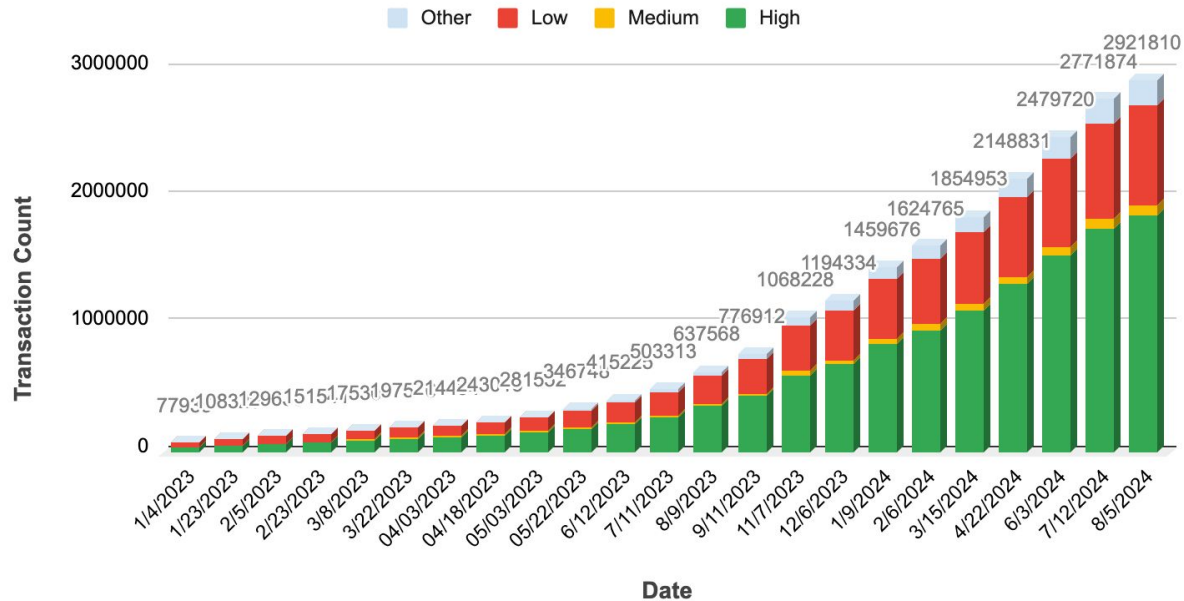
Confidence quarterly breakdown

19

What's coming next

Default rules

- Global rule set used when confidence is less than X%
- Results in better predictions & yields more time saved

Train Your Model

- Supporting additional dimensions
- Auto-splitting
- Leverage automation with clients without GL history
- Enhanced control - can specify descriptions, amounts
- Workflows - adding assignees, marking transactions needs review, excluding transactions

ML Improvements

- Preprocessing
- Named entity recognition models
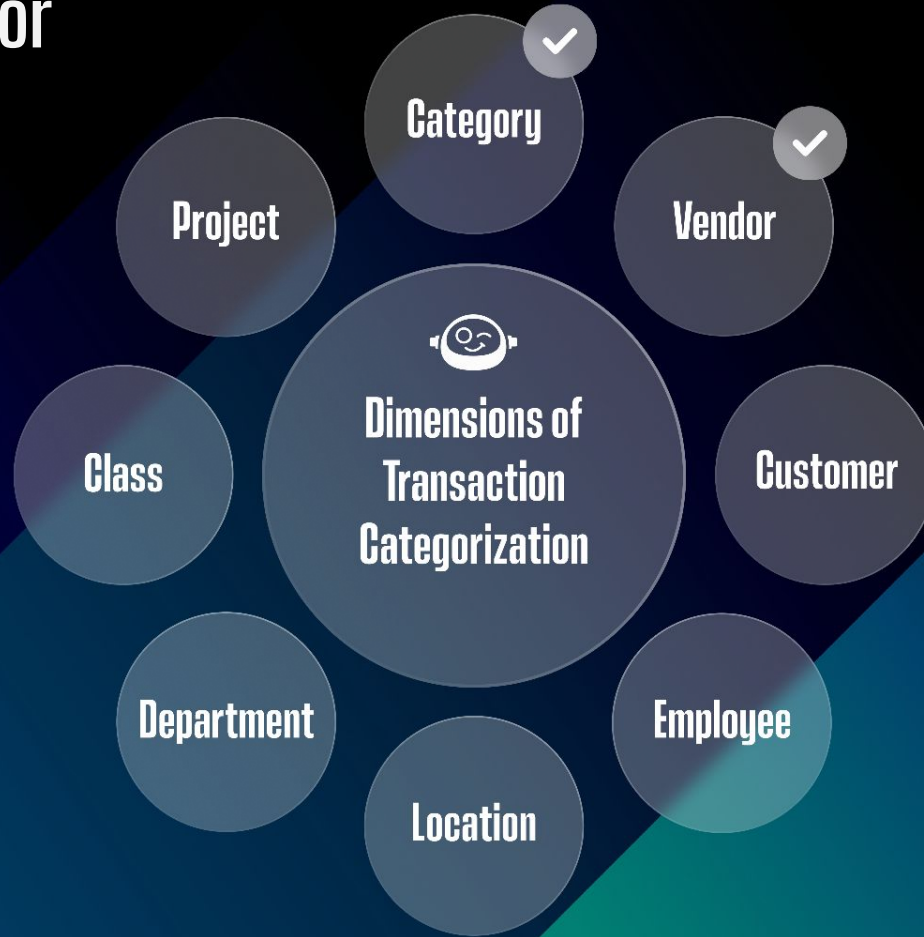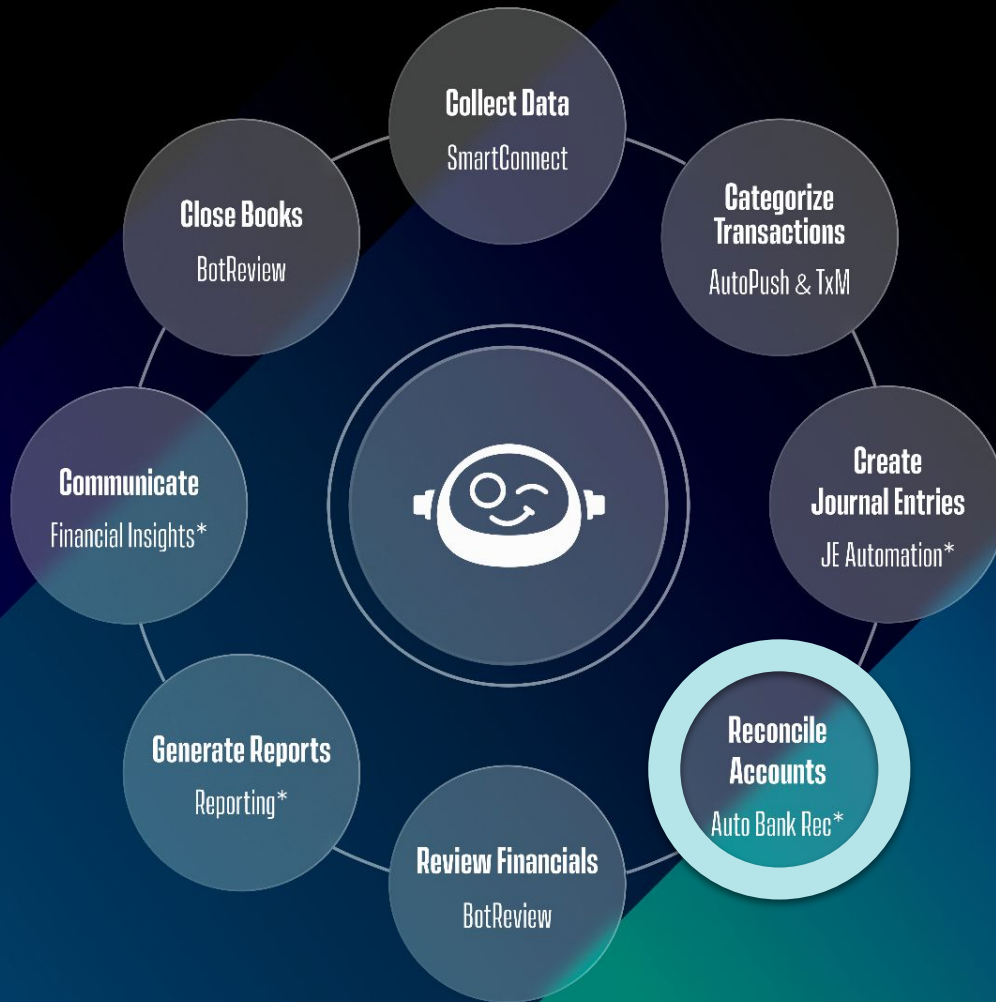
MODEL IMPROVEMENTS

PREPROCESSING

Transaction descriptions often contain noise such as irrelevant characters, variations in formats, and unnecessary details, which can confuse the model.

NAMED ENTITY RECOGNITION

More focus on meaningful information and discarding irrelevant information in the transaction description.

Have you ever had to review a bank statement line-by-line, trying to figure out what's causing the difference in your reconciliation?

A Glance at Auto Bank Rec

Auto Bank Rec automates reconciliations with bank feeds and statements to identify outliers with smart, actionable solutions.
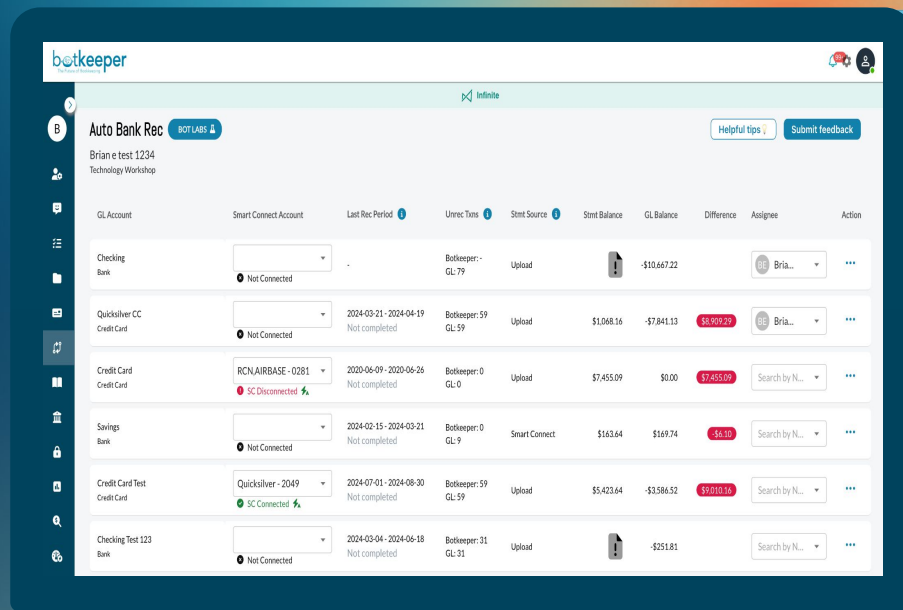
Opportunities

COLLECTION

Leverage technology to collect the transactions and statements

COMPARISON

Leverage technology to compare the data sets and determine outliers

COLLABORATION

Streamline collaboration around reconciling accounts

![AI UNCHAINED — Accounting Intelligence. No boundaries. No limits.]

Real-time bank feed reconciliation also available!

Automated Bank Reconciliation

Retrieve Bank Statements
Automatically retrieved for supported institutions or uploaded via Documents

Extract Data
OCR technology is used to parse account and transactions details

Botkeeper Data Review (HITL)
Extracted data is reviewed by Botkeeper experts

Transactions Automatically Matched
Statement transactions are programmatically matched to transactions in the GL

Reconciliation Finalized
Simple UI allows users to review and complete reconciliation

27

DATA EXTRACTION FROM BANK STATEMENTS

RULE BASED TEMPLATES

- Pre-defined templates that specify the exact location and format of the data to be extracted
- Looks for specific keywords or locations in a structured document

MACHINE LEARNING MODELS

- Models that can be trained to recognize the layout and structure of bank statements
- Makes use of OCR and natural language processing (NLP) techniques to recognize and extract data

TEMPLATES ➕ ML MODELS

TRAINING DATA

Don't require a large amount of labeled data for training and are quick to deploy

SCALABILITY

Can handle a wide variety of institutions and layouts with format variations

ACCURACY

Accuracy is very high and the results are consistent

ABILITY TO LEARN

Can improve over time with more data and fine-tuning

HUMAN IN THE LOOP

- Combines the efficiency of automation with the precision of human oversight to ensure high accuracy
- Continuous feedback loop enhances system learning, allowing automation to become more accurate over time

AUTOMATED MATCHING

- Ensures financial accuracy by automatically comparing and aligning data between internal records and bank statements
- Reduces the need for manual intervention and speeds up the reconciliation process

33%
Completely automated reconciliations

88%
Automatically reconciled transactions

What's coming next

Custom models for complex statements

◇ Combined statements
  · Multiple bank accounts
  · Credit cards with subaccounts

Supporting Subaccounts

◇ One reconciliation including main & subaccounts

Bulk Actions

◇ Adding/Editing/Deleting
◇ Categorizing statement transactions

Have you ever dreaded a client meeting due to not knowing what to talk about?

34

A Glance at
Financial Insights

Financial Insights combines data and GenAI to tell the story of how the business is doing, what is the driving the performance and areas of opportunity; allowing firms to offer valuable advisory services and scale their practice.
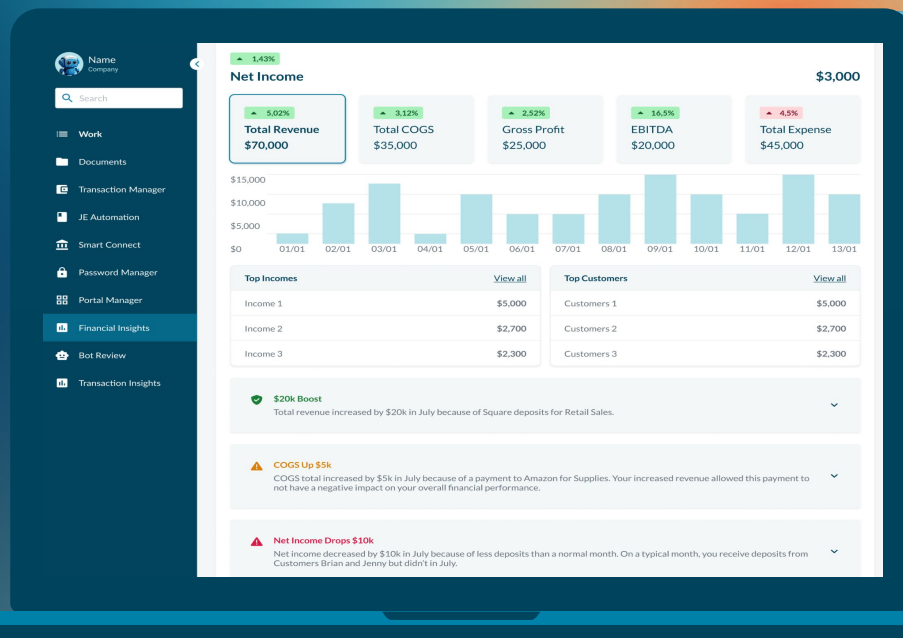
Opportunities

COLLECTION

Leverage technology to collect the financial data

COMMUNICATE

Take advantage of GenAI to tell the story in an easy-to-understand format

CLIENT VALUE

Set your firm apart from the rest by providing your clients with actionable insights

Botkeeper's Financial Insights

CAS FOCUSED
Let's get the advisory conversations started!

INCOME STATEMENTS
We'll be starting with a history of income statements, driven by your client's general ledger data.

TRENDS & OPPORTUNITIES
We'll layer statistical analysis and trending with...

DYNAMIC MESSAGING
Large language models to provide comprehensive and easy to understand insights.

STATISTICAL MODELS

Mathematical techniques to perform quantitative analysis that are highly accurate and grounded in well-established methodologies

LARGE LANGUAGE MODELS

Interpret results, providing context, summarizing key points, and translating technical insights into easy-to-read narratives

FINANCIAL INSIGHTS

Quantitative accuracy and qualitative explanation, creating a comprehensive and easy-to-understand financial analysis

THANKS!

Do you have any questions?

jblackwell@botkeeper.com

bemposimato@botkeeper.com

nvimalesh@botkeeper.com

botkeeper.com

AI UNCHAINED
Accounting Intelligence. No boundaries. No limits.