TAXXCISE

AI-Powered Tax & Accounting Services Platform

Complete End-to-End Technical Documentation

Built by Muhammad Ibrahim

December 2024

Executive Summary

Taxxcise is a comprehensive AI-powered lead management and automation platform specifically designed for CPA practices. The system combines a professional static website with sophisticated backend automation, enabling small independent accounting practices to compete with larger firms through intelligent automation and professional web presence—all at zero monthly cost.

Project Vision

This project was built to provide a low-cost, high-value solution for an independent CPA. The goal was to demonstrate that sophisticated automation platforms can be built cost-effectively using free tier cloud resources when properly architected, while delivering enterprise-grade functionality.

Key Achievements

- **Zero Monthly Cost:** Entire platform runs within GCP free tier limits
- **AI-Powered Lead Processing:** Automatic lead qualification and personalized responses using Google Gemini
- **Multi-Layer Security:** reCAPTCHA, honeypot, rate limiting, Cloud Function proxy, Secret Manager
- **Automated Workflows:** Contact form processing, email responses, monthly newsletters
- **Privacy-First Analytics:** Self-hosted Umami with GDPR compliance
- **CI/CD Pipeline:** Automated deployment via GitHub Actions

System Architecture

The Taxxcise platform employs a modern containerized architecture running on Google Cloud Platform. All services are orchestrated using Docker Compose, with Caddy serving as the reverse proxy and handling automatic HTTPS certificate provisioning.

Infrastructure Overview

Component	Specification
Cloud Provider	Google Cloud Platform (GCP)
Instance Type	e2-micro (1 vCPU, 1GB RAM, 30GB Disk)
Operating System	Ubuntu 22.04 LTS
Container Runtime	Docker with Docker Compose
Reverse Proxy	Caddy (Auto HTTPS via Let's Encrypt)
Database	PostgreSQL 15 (shared between n8n and Umami)
DNS Provider	DuckDNS (Dynamic DNS)

Container Services

The platform runs six containerized services within a Docker bridge network:

Service	Port	Purpose
Caddy	80, 443 (published)	Reverse proxy, SSL termination, static file serving
n8n	5678 (internal)	Workflow automation engine, webhook receiver
PostgreSQL	5432 (internal)	Shared database for n8n and Umami
Umami	3000 (internal)	Privacy-first web analytics
Watchtower	N/A	Automatic Docker container updates (daily at 4 AM)

Network Topology

All containers communicate via a Docker bridge network (n8n-network). External traffic enters through Caddy, which terminates SSL and routes requests to the appropriate internal service. The PostgreSQL database is not exposed to the internet and is only accessible from within the Docker network.

Website & Frontend

The Taxxcise website is a modern, mobile-first static site built with HTML5, CSS3, and vanilla JavaScript. The design emphasizes professionalism and trust, critical factors for a CPA practice.

Design System

- **Primary Color:** Navy (#1B2A41) - Conveys professionalism and trust
- **Accent Color:** Gold (#C5A059) - Adds warmth and prestige
- **Typography:** System fonts for optimal performance
- **Layout:** CSS Grid and Flexbox for responsive design

Page Structure

Page	Purpose
index.html	Homepage with hero section, services overview, and call-to-action
about.html	Company background, credentials, and team information
services.html	Detailed service offerings with pricing indicators
contact.html	Contact form with multi-layer security protection
privacy.html	Privacy policy and data handling practices
disclaimer.html	Legal disclaimers and terms of service
record-retention.html	IRS record retention guidelines for clients
state-tax-forms.html	Links to state tax form resources
thank-you.html	Confirmation page after form submission

Security Architecture

Security is implemented as a multi-layered defense system, protecting against various attack vectors while maintaining a seamless user experience.

Contact Form Security Layers

The contact form implements six distinct security measures:

1. Google reCAPTCHA v2

Checkbox-based CAPTCHA verification prevents automated bot submissions. The token is validated server-side through Google's verification API before any form data is processed.

2. Honeypot Field

A hidden form field (labeled 'company') is invisible to real users but automatically filled by bots. Any submission with this field populated is silently discarded, appearing successful to the bot.

3. Google Cloud Function Proxy

The actual n8n webhook URL is never exposed to the browser. Instead, form submissions are sent to a Cloud Function that retrieves the real webhook URL from Secret Manager and forwards the request. This prevents attackers from directly targeting the webhook endpoint.

4. Rate Limiting

The Cloud Function implements IP-based rate limiting of 5 requests per 60 seconds. This prevents abuse while allowing legitimate users to retry failed submissions.

5. CORS Protection

The Cloud Function only accepts requests from the legitimate domain (taxxcise.duckdns.org). Cross-origin requests from other domains are rejected.

6. Server-Side Validation

All form fields are validated on the server for required values, proper formatting (especially email addresses), and reasonable length limits.

Infrastructure Security

- **HTTPS Enforcement:** All traffic encrypted via Let's Encrypt certificates
- **SSH Key Authentication:** Password-based login disabled
- **GCP Firewall:** Only ports 80, 443, and 22 open
- **Container Isolation:** Services run in isolated Docker containers
- **Read-Only Volumes:** Website files mounted as read-only
- **Security Headers:** HSTS, X-Frame-Options, CSP, X-Content-Type-Options

Automation Workflows

The platform uses n8n, a self-hosted workflow automation tool, to orchestrate all automated processes. Four primary workflows handle lead capture, email communication, and newsletter distribution.

Workflow 1: Contact Form Response & Enrichment

Trigger: Webhook from contact form submission

This workflow captures incoming leads, enriches them with AI analysis, and stores them in the CRM.

Processing Steps:

1. Receive form data via secure webhook endpoint
2. Validate reCAPTCHA token with Google's verification API
3. Extract metadata: IP address, timestamp, user agent
4. Send data to Google Gemini for AI enrichment
5. Store enriched lead in Notion CRM database
6. Return success response to browser

AI Enrichment Analysis:

- **Likely Profession/Role:** Inferred from name, email domain, and message content
- **Inferred Pain Point:** Primary concern identified from message context
- **Lead Temperature:** Cold, Warm, or Hot based on urgency indicators
- **Suggested Opening Line:** Personalized conversation starter for follow-up

Workflow 2: Send Email (Automated Response)

Trigger: New entry in Notion database

Sends personalized follow-up emails to new contacts using AI-generated content based on lead enrichment data.

Processing Steps:

1. Monitor Notion database for new contact entries
2. Retrieve contact details and AI enrichment data
3. Generate personalized email using Google Gemini
4. Format email as HTML with professional template
5. Send via Gmail API with proper headers
6. Update Notion record with 'Email Sent' status

Workflow 3: Monthly Newsletter

Trigger: Scheduled (1st of each month at 9 AM)

Aggregates tax news and distributes personalized newsletters to opted-in subscribers.

Content Sources:

- IRS news releases for the current month
- Tax Foundation RSS feed articles
- Upcoming quarterly tax deadlines and reminders

Workflow 4: Newsletter Unsubscribe

Trigger: Webhook from unsubscribe link

Handles unsubscribe requests by locating the subscriber in Notion and updating their newsletter_opt_in status to false.

CRM Integration

The platform uses Notion as a lightweight CRM, storing all contact information and lead enrichment data in a structured database.

Contact Database Schema

Field	Type	Description
Name	String	Contact's full name
Email	String	Contact email address
Message	String	Original form submission message
IP Address	String	Submitter's IP for geolocation
Likely Profession/Role	String	AI-inferred profession
Inferred Pain Point	String	AI-identified primary concern
Lead Temperature	Select	Cold, Warm, or Hot classification
Suggested Opening Line	String	AI-generated personalized opener
Lead Status	Select	New, Contacted, Converted, etc.
Newsletter_opt_in	Boolean	Newsletter subscription status
submitted_at	DateTime	Form submission timestamp
updated_at	DateTime	Last record update

Analytics & Monitoring

Umami Analytics

The platform uses Umami, a self-hosted privacy-first analytics solution, to track website usage without requiring cookies or compromising user privacy.

Key Features:

- **No Cookies Required:** Compliant with GDPR without cookie consent banners
- **Self-Hosted:** All data stays on your infrastructure
- **Real-Time Dashboard:** Live visitor statistics and behavior tracking
- **Custom Event Tracking:** Track form interactions, button clicks, and conversions

Tracked Custom Events:

- Contact form starts and completions
- Service card clicks
- Navigation link interactions
- CTA button clicks
- Newsletter subscription conversions

Infrastructure Monitoring

GCP Cloud Monitoring provides visibility into infrastructure health:

- CPU and memory utilization
- Disk I/O and storage usage
- Network throughput
- Docker container health status

DevOps & CI/CD

Automated Deployment Pipeline

GitHub Actions powers the CI/CD pipeline, enabling automatic deployment on every push to the main branch.

Deployment Process:

1. Developer pushes code to GitHub main branch
2. GitHub Actions workflow triggers automatically
3. Workflow establishes SSH connection to GCP server
4. Server pulls latest code from GitHub repository
5. Website is live within 30 seconds

Automatic Container Updates

Watchtower runs daily at 4 AM, automatically pulling and deploying updated Docker images. This ensures all services stay current with security patches and feature updates without manual intervention.

Backup Strategy

Automated daily backups include:

- PostgreSQL database dump (n8n workflows and Umami analytics)
- n8n data volume (credentials and execution history)
- Caddy SSL certificates and configuration
- 14-day retention with automatic cleanup of older backups

Cost Analysis

Current Monthly Cost: $0

The entire platform operates within GCP's free tier limits, resulting in zero monthly infrastructure costs.

Free Tier Resource Allocation

Resource	Free Tier Limit	Actual Usage
e2-micro Instance	744 hours/month	744 hours (100%)
Standard Disk	30 GB	~8 GB (27%)
Static IP (attached)	Free	Free
Cloud Functions	2M invocations	~100/month (<1%)
Secret Manager	6 active secrets	2 secrets (33%)
Network Egress	1 GB/month	~0.5 GB (50%)

Scaling Cost Projections

If the platform needs to scale beyond free tier limits, here are the projected costs:

Upgrade	Monthly Cost	When Needed
e2-small (2GB RAM)	~$14	More than 5 concurrent users
e2-medium (4GB RAM)	~$27	More than 10 concurrent users
Cloud Armor (DDoS)	~$15	Under frequent attack
Managed PostgreSQL	~$10	Database exceeds 5GB

Complete Technology Stack

Frontend

- HTML5/CSS3 with modern semantic markup
- Vanilla JavaScript (no frameworks)
- Mobile-first responsive design

Backend & Infrastructure

- Google Cloud Platform (GCP e2-micro)
- Docker & Docker Compose for containerization
- Caddy for reverse proxy and auto-HTTPS
- PostgreSQL 15 for persistent storage
- Watchtower for automatic container updates

Automation & AI

- n8n for workflow automation
- Google Gemini Pro for AI lead enrichment and email generation
- Google Cloud Functions for serverless proxy

Integrations

- Gmail API for email delivery
- Notion API for CRM functionality
- Google reCAPTCHA v2 for bot protection
- Secret Manager for credential storage

Analytics & Monitoring

- Umami Analytics (self-hosted, privacy-first)
- GCP Cloud Monitoring for infrastructure
- n8n execution logs for workflow debugging

DevOps

- GitHub for version control
- GitHub Actions for CI/CD
- DuckDNS for dynamic DNS

Future Enhancement Possibilities

While the current implementation is complete and production-ready, several enhancements have been documented for potential future development:

Vapi.ai Voice Assistant Integration

- AI-powered phone receptionist for client inquiries
- Automated appointment scheduling via voice
- Document collection reminders through outbound calls
- Voice-to-text transcription integration with n8n workflows

Accounting Software Integration

- QuickBooks Online/Desktop sync for client financial data
- Xero integration for bookkeeping automation
- Bi-directional sync for invoices and payments

Frontend Improvements

- Interactive tax estimation calculators
- WCAG 2.1 AA accessibility compliance
- Dark mode option
- Animated micro-interactions

Conclusion

The Taxxcise platform demonstrates that sophisticated, enterprise-grade automation can be achieved at zero cost through careful architectural decisions and strategic use of cloud free tiers. The project combines modern web development, containerization, AI integration, and workflow automation into a cohesive system that serves real business needs.

Key Takeaways

1. **Strategic Architecture:** Thorough upfront research on self-hosting, free tier limitations, and technology trade-offs enabled cost-effective implementation.
2. **AI Integration:** Google Gemini enables intelligent lead processing that would otherwise require manual effort or expensive third-party services.
3. **Security First:** Multi-layer security implementation protects both the business and its clients.
4. **Automation:** End-to-end automation eliminates manual processes and ensures consistent client communication.
5. **Scalability:** The architecture can scale with growing business needs while maintaining cost efficiency.

— END OF DOCUMENTATION —

Built by Muhammad Ibrahim | December 2024