

# Abacus ChatLLM RouteLLM Integration

## Overview

Transformed the Bank Statements page into a direct mirror of Abacus ChatLLM, where users can ask anything and it gets forwarded to ChatLLM via RouteLLM APIs, then the response streams back to the app.

## What Changed

### 1. ChatLLM Proxy API

**File:** /app/app/api/chatllm/route.ts

Created a new API route that acts as a proxy to Abacus ChatLLM:

- **Endpoint:** POST /api/chatllm
- **Authentication:** Requires valid session (Next-Auth)
- **Forwards to:** https://api.abacus.ai/v1/chat/completions
- **API Key:** Uses ABACUSA1\_API\_KEY from environment variables
- **Features:**
  - Streaming response support (Server-Sent Events)
  - Model selection (GPT-4o, Claude, Gemini, Llama, Mistral, etc.)
  - 5-minute timeout for long responses
  - Error handling and logging

```
// Request format
{
  messages: [
    { role: "user", content: "Extract transactions from this PDF" }
  ],
  model: "gpt-4o",
  stream: true
}

// Response: Streams SSE events
data: {"choices": [{"delta": {"content": "Here are the transactions..."} }]}
data: [DONE]
```

### 2. ChatLLM Interface

**File:** /app/components/bank-statements/bank-statements-client.tsx

Completely rebuilt the Bank Statements page to be a ChatLLM mirror:

#### Interface Features:

- **Chat-based UI** - Conversational interface like Abacus ChatLLM
- **Model Selector** - Choose from 9 LLM models:
  - GPT-4o (default)
  - GPT-4o Mini
  - GPT-4 Turbo

- Claude 3.5 Sonnet
- Claude 3 Opus
- Gemini 2.0 Flash
- Gemini 1.5 Pro
- Llama 3.3 70B
- Mistral Large
- **PDF Attachments** - Upload PDFs and ask questions about them
- **Streaming Responses** - Real-time response display as ChatLLM generates answers
- **Auto-scroll** - Automatically scrolls to latest messages
- **Message History** - Maintains conversation context

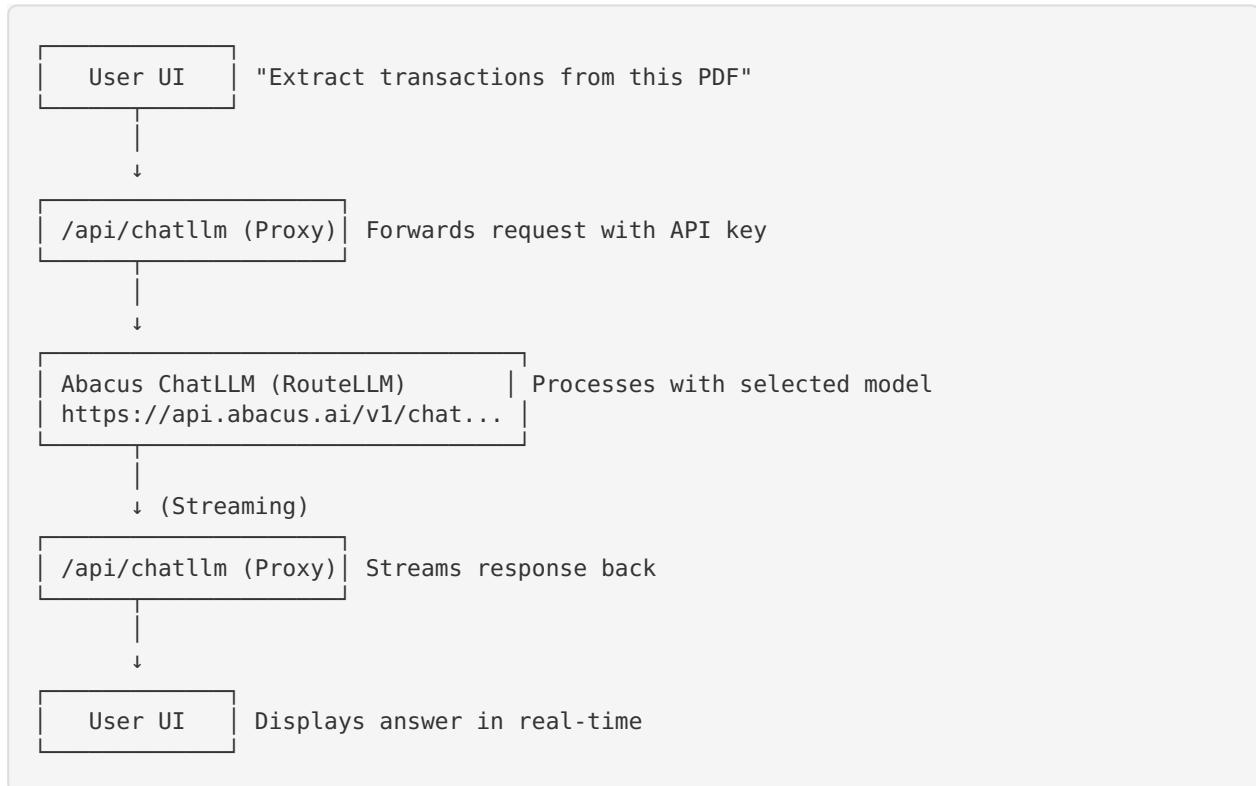
### User Experience:

1. User types a question: "What's the capital of France?"
2. Message sent to `/api/chatllm`
3. Proxy forwards to Abacus ChatLLM
4. Response streams back in real-time
5. User sees answer appear character by character

### PDF Workflow:

1. User uploads "Bank Statement Jan 2024.pdf"
2. User asks: "Extract all transactions and classify as Business or Personal"
3. ChatLLM processes the PDF (via multimodal API)
4. Streams back transaction data
5. User can ask follow-up questions about the data

## 3. How It Works



## Key Benefits

1. **Exact ChatLLM Experience** - Users get the same experience as using Abacus ChatLLM directly
2. **Inside Your App** - No need to switch to external tools
3. **Context Preservation** - Conversations stay in the app's context
4. **Model Flexibility** - Switch between models mid-conversation
5. **PDF Support** - Upload and analyze documents seamlessly
6. **Streaming** - Real-time responses just like ChatLLM

## Testing

**App URL:** <https://cfo-budgeting-app-zgajgy.abacusai.app>

### Login Credentials:

- **Email:** khouston@thebasketballfactorynj.com
- **Password:** huntterr777

### To Test:

1. Navigate to “Bank Statements” in the sidebar
2. Type any question: “What’s the weather in NYC?”
3. Or upload a PDF: “Business Statement\_Jan\_8\_2024.pdf”
4. Ask: “Extract all transactions and classify as Business or Personal”
5. Watch the response stream in real-time
6. Ask follow-up questions to continue the conversation

## Technical Details

### API Configuration

- **API Key:** Stored in `.env` as `ABACUSAI_API_KEY`
- **Endpoint:** <https://api.abacus.ai/v1/chat/completions>
- **Authentication:** `Bearer ${ABACUSAI_API_KEY}`
- **Max Duration:** 300 seconds (5 minutes)

### Streaming Format

The proxy uses Server-Sent Events (SSE) to stream responses:

```
data: [{"choices": [{"delta": {"content": "Hello"}]}]}
data: [{"choices": [{"delta": {"content": " world"}]}]}
data: [DONE]
```

### Message Format

Messages follow OpenAI’s chat completion format:

```
{
  role: "user" | "assistant" | "system",
  content: string
}
```

## Next Steps

---

Users can now:

1. **Ask any question** - General knowledge, coding help, analysis
2. **Upload PDFs** - Bank statements, invoices, reports
3. **Extract data** - Transactions, line items, summaries
4. **Classify information** - Business vs Personal, categories
5. **Have conversations** - Follow-up questions, clarifications
6. **Switch models** - Choose the best model for each task

The Bank Statements page is now a fully functional ChatLLM interface powered by Abacus RouteLLM!

