**Unstract Platform – Overview & Project Setup Documentation**

**1. Platform Overview :**

Unstract is an open-source, no-code platform that automates complex business processes involving long, unstructured documents, with human-in-the-loop support.

Key Capabilities

* Prompt Studio for custom document types using LLMs, Vector DBs, and embedding models.
* Source & Destination Integration with object storage, file systems, and databases.
* Workflow Deployment as APIs, ETL pipelines, or Q&A apps.

**2. Editions :**

**1)Cloud Edition**

* **Description:** Fully managed, hosted; zero setup required.
* **Trial:** 14-day free trial (no credit card).
* **Enterprise Features:**
  + **LLMChallenge:** Uses two LLMs for high accuracy (right answer or no answer).
  + **SinglePass Extraction:** Up to 8× lower token usage.
  + **SummarizedExtraction:** Up to 6× lower token usage.
  + **Human Quality Review:** Side-by-side view of extracted vs. source data.
  + **SSO Support:** Multiple authentication options.
* **Security & Compliance:** SOC 2 Type II, HIPAA, GDPR, ISO 27001 certified.

**2)Open Source Edition**

* **Description:** Self-managed for technical users.
* **Missing Features:** No LLMChallenge, SinglePass, SummarizedExtraction, Human Quality Review, or SSO.
* **System Requirements:** 8GB RAM, Linux/MacOS, Docker, Docker Compose, Git.
* **Setup Steps:**
  1. Download release or clone repository.
  2. Run ./run-platform.sh.
  3. Access at http://frontend.unstract.localhost (login: unstract/unstract).

**3)On-Premise Edition**

* **Description:** Self-hosted for industries with strict data regulations.
* **Infrastructure:** Runs on Kubernetes; supports AWS, Azure, GCP.
* **Features:** Includes all Cloud Edition enterprise features.
* **Advantages:** Cost savings, reduced latency, high accuracy.

**3. Steps to Create a Project in Prompt Studio**

**Step 1: Create New Project**

* Sign in to Unstract.
* Go to **Prompt Studio** → **New Project**.
* Fill in:
* Tool Name: *Credit Card Statements Parser*
* Author/Org Name: *Your Name*
* Description: *Parses US Credit Card Statements*
* Icon: *Credit Card*
* Click **Save** → Open Project.

**Step 2: Understand Prompt Studio Layout**

* **Prompt Engineering Pane:** Add/manage prompts for data extraction.
* **Document Pane:** View original/raw/summary formats.
* **Document Manager:** Manage uploaded documents.
* **Project Settings:** LLM profiles & configurations.
* **Export Project:** Save for workflow or API.

**Step 3: Configure Project Settings**

* **A. Create LLM Profile**
* Name: *OpenAI GPT 3.5 + Qdrant + Ada + LLMWhisperer*
* LLM: *OpenAI*
* Vector DB: *Qdrant*
* Chunk Size & Overlap: 0
* Embedding Model: *OpenAI embedding model*
* Text Extractor: *LLMWhisperer*
* Click **Add**.
* **B. Add Preamble**
* Guidance: Extract & summarize accurately using provided context.
* Save.
* **C. Add Postamble**
* Instruction: Only return extracted info without explanations.
* Save.

**Step 4: Add Documents**

* Open **Documents Manager** → Upload PDFs.
* Files initially show *Not Indexed / Not Summarized* (updates after running prompts).

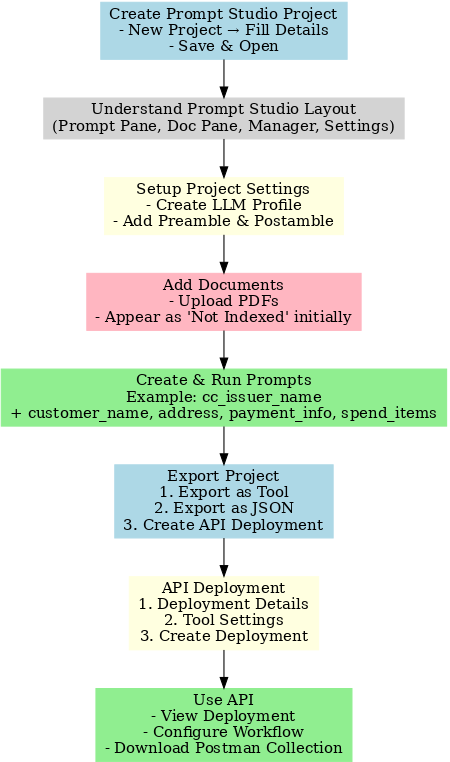
**Step 5: Create & Run Prompts**

* Example: **Credit Card Issuer Name**
* Field Name: cc\_issuer\_name
* Field Prompt: *Extract primary brand name of the credit card issuer.*
* Click **Run** → Check coverage info.
* **Additional Prompts:**
* **Customer Name:** Extract full name (capitalized).
* **Customer Address:** JSON { full\_address, city, zip }
* **Payment Info:** JSON { new\_balance, minimum\_payment, due\_date }
* **Spend Line Items:** JSON array { spend\_date, spend\_description, amount}

**Step 6: Export Project**

* **As Tool:** For Workflow Studio.
* **As JSON:** For sharing/backups.
* **API Deployment:** Deploy directly as an API.

**Step 7: Create API Deployment**

* **Export Options:** → *Create API Deployment*
* **Details:**
* Display Name: *Credit Card Parser API*
* Description: *API for parsing credit card statements*
* API Name: Unique identifier.
* **Tool Settings:**
* Config Name: credit\_card\_parser
* Challenger LLM: Select for LLMChallenge
* Enable features: LLMChallenge, Highlight, SummarizedExtraction, SinglePass Extraction.
* **Deployment Success:**
* View API details.
* Download Postman Collection for testing.
  1. **Flowchart**

**5.Pricing Overview**

**Monthly Plans**

* Starter – $499/month, includes 5,000 pages/month, overage $0.10/page
* Growth – $2,249/month, includes 25,000 pages/month, overage $0.09/page

**Annual Plans *(Billed annually)***

* Starter – $416/month (equivalent), includes 60,000 pages/year, overage $0.10/page
* Growth – $1,874/month (equivalent), includes 300,000 pages/year, overage $0.09/page

**Enterprise** – Custom pricing available upon request.