SYSTEM DESIGN

Contents

1. Introduction

1.1. Purpose

The purpose of this document is to define the system architecture and design for the CartographAI, including its backend, persistence store, Rule library, Co-Pilot, AI engine, and CartographAI Web Application. It describes the structural components, interactions, and communication protocols to ensure the system meets its functional, non-functional, and compliance requirements

1.2. Scope

This document outlines the architecture and design for the entire CartographAI system, covering:

- Software: Rule library AI engine, co-pilot, backend, and CartographAI Web Application.
- Persistence Store: Secure storage and retrieval of data.
- Communication Protocols: Secure APIs for data exchange.

1.3. Intended Audience

- Software Developers
- IT Security Officers
- Quality Assurance Teams
- Project Managers

2. System Overview

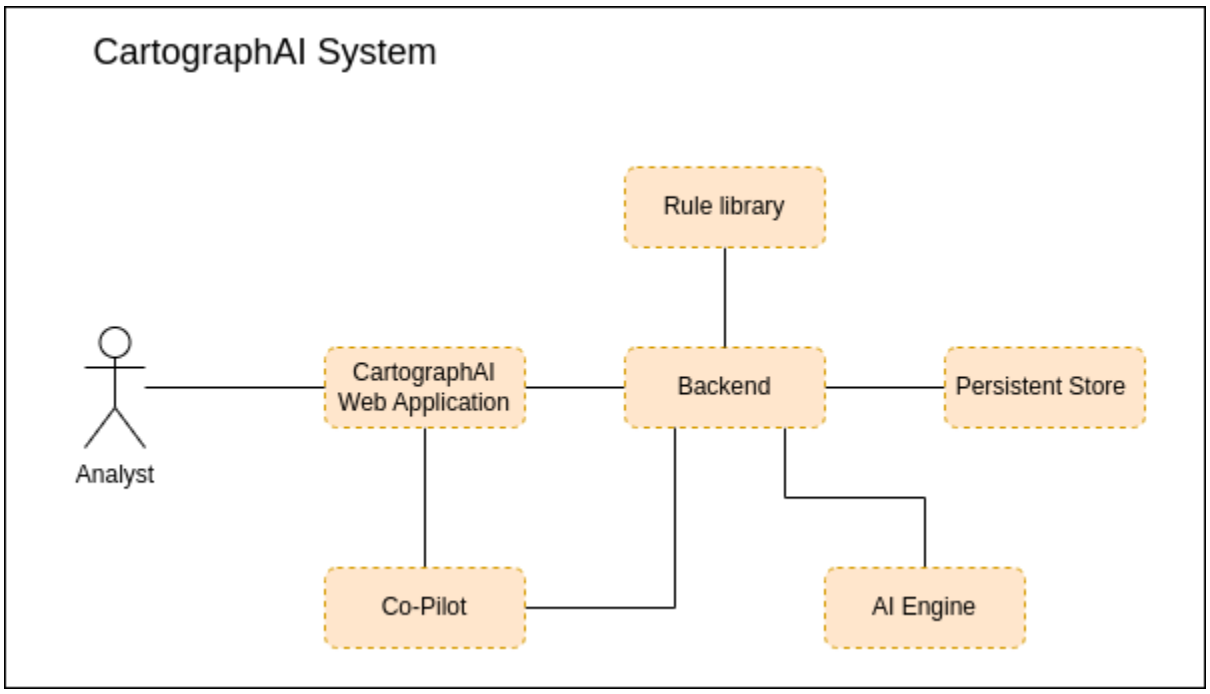
CartographAI is an automated data analysis, mapping, and transformation engine powered by AI that can accurately deliver these responsibilities in hours rather than months. Its primary components
include:

- CartographAI Web Application: Web-based interface for Analysts to discover, Map, transform, and move data from source system to destination system.
- Rule library: Provides mapping rules for all significant Financial Crime Compliance applications, beginning with the most prevalent.
- Co-Pilot: The Discover Co-Pilot is an AI-powered virtual assistant. Users can interact with the co-pilot to clarify Discover insights and drill deeper into source databases, tables, attributes, and interrelationships.
- Backend: Provides centralized processing, business logic, and API services.

- Persistence Storage: Stores discovered assets, mapping details, Transformation logic, and audit logs securely.
- AI Engine: Provides core functionality of discovery, mapping, transformation, and co-pilot.

3. System Architecture

3.1. High-Level System Architecture



- CartographAI Web Application
  - Accesses and displays discovery assets, mapping, and transformation data by communicating with the backend via secure REST APIs.
  - Retrieves transformation logic, generates audit report, and data lineage, and executes commands to generate discovery assets, mapping, and transformation through backend interaction.
- Co-Pilot
  - Accesses discovery assets by communicating with the backend via secure REST APIs
  - Retrieves discovery knowledge base through backend interaction.
- Rule library
  - The Map module of CartographAI leverages the Discover module output along with a knowledge base of mapping rules to propose source-to-target mappings for any target FCC application.
- Persistence store

    - Stores all assets generated by CartograpAI
    - Functions as the dedicated storage layer, accessible solely through backend APIs.
    - Restricts direct access to ensure data security and integrity.
- AI Engine
    - It serves as the core engine for discovery, mapping, transformation, and co-pilot
    - The backend interacts with the AI engine to process user requests from the Web Application
- Backend
    - It serves as the central intermediary connecting the Web Application, rule library, AI engine, Co-Pilot, and database.
    - Handles business logic, enforces security measures, encrypts data, and manages storage operations.
    - Interacts with the persistence store for storing and retrieving data

4. Component Overview

#	Component	Role
1	CartographAI Web Application	Web-based interface for Analysts to discover, Map, transform, and move data from source system to destination system.
2	Co-Pilot	The Discover Co-Pilot is an AI-powered virtual assistant.
3	Rule Library	Provides mapping rules for all significant Financial Crime Compliance applications, beginning with the most prevalent.
4	Backend	Provides centralized processing, business logic, and API services.
5	AI Engine	Provides core functionality of discovery, mapping, transformation, and co-pilot.
6	Persistent Store	Stores discovered assets, mapping details, Transformation logic, and audit logs securely.

5. Detailed design

5.1. Software Design

Component	Description	Key Features
CartographAI Web Application	Web-based interface for discovery, mapping, transformation, and co-pilot	User-friendly navigation, co-pilot

Backend	Centralized processing and data management	Secure API, data encryption, role-based access control
Co-Pilot	AI BOT to interact with generated assets.	User friendly chatbot

6. Communication Protocol

#	Protocol	Usage
	REST API	Facilitates communication between UI application, backend, and CartographAI Web Application.

7. Security Measures

- Data Encryption: All data in transit and at rest is encrypted using AES-256.
- Authentication: Multi-Factor Authentication (MFA) and Role-Based Access Control
- (RBAC).

8. Error Handling and Logging

Component	**Error Handling**	**Logging**
CartographAI Web Application	Displays user-friendly error messages and recovery options	Maintains logs of user interactions and report generations.
Backend	Handles API failures gracefully and retries failed transactions.	Tracks API usage and failed queries.
Co-Pilot	Notifies users of missing or invalid data.	Maintains logs of user interactions and report generations

9. Testing and Validation

9.1. Unit Testing

- Testing of individual modules (e.g., Co-Pilot, user authentication).

9.2. Integration Testing

- Verification of communication between UI, backend, Rule library, AI Engine, Co-Pilot, and database.

9.3. System Testing

- End-to-end testing of the entire system, including real-world scenarios.

9.4. User Acceptance Testing (UAT)

- Validation by clinical users in simulated environments.

10. System Requirements

ID	Name	Description	Linked BRD requirement
SDD-1	User authentication	The CartographAI Web Application shall require users to authenticate using secure credentials before accessing any features.	BRD-50, BRD-51
SDD-2	User authentication	Configurable authentication. It should integrate with local db or LDAP or Microsoft AD or any other Identity Provider using a standard JWT token. Should support SSO.	BRD-50
SDD-3	Role based access control	Configurable role management	BRD-51
SDD-4	Project Management	Create, update, and delete projects. Allow management of project access.	??
SDD-5	Configure	Allow Configuration of the source	BRD-1, 20

		system, target system, rule library module, or any third party system. Allow validation of the configuration information. Allow Import/ export of configuration between systems.	
SDD-6	Discover - RDBMS	Use AI engine to discover source system. Specific AI engines should be configurable. Allow configurable AI engines public or private.	BRD-8,9,10, 11,12
SDD-7	Discover - Assets	Generate assets using AI engine and store all the details using a markdown language. Allow download of the assets as PDF.	BRD-13,14, 15,16,17, 19
SDD-8	Co-Pilot	Allow users to interact & query from the discovered assets. Allow users to create/ manage their own AI prompt. Allow the user to share the prompts.	BRD-18
SDD-9	Map - Configuration	Configure and manage rules library	BRD-21
SDD-10	Map - Source to target mapping	Allow input and target system mapping based on the rules defined in the library. The system should provide a confidence level of mapping. Users should have the option to edit the mapping and provide a feedback loop to the system for continuous improvement. Allow export/ import of the mapped details.	BRD-22,27, 28, 29, 30,31, 32
SDD-10	Map - visualization	Manage the mapped details between the source and target. Provide a visual tree of the mapping	BRD-32
SDD-11	Transform - Configure	Configure and manage the transformation output required e.g. SQL, Python, etc.	??
SDD-12	Transform - Manage	Generate ETL logic based on configuration and save the details. It should provide a confidence level. Allow the user to edit the generated ETL logic. Allow the user to simulate	BRD-34, 36

		the ETL.	
SDD-13	Transform - Execute	Execute the ETL logic. Run validation test. Generate execution report. Download the execution report as PDF document	BRD-36
SDD-13	Governance - Record	Allow record of all activities in the system	BRD-48
SDD-14	Governance - Traceability	Visual display of source to target lineage	BRD-48
SDD-15	Security - Logging	Appropriate logging is to be maintained. Allow scheduled backup of log. Allow configuration of managing logs	BRD-52
SDD-16	Application logging	Allow configurable application login.	BRD-52
SDD-17	Licensing	Allow managing license key by the CartographAI team. Allow to generate, renew, revoke license key	BRD-56
SDD-18	AI Engine	Allow configurable AI engine. OpenAI/ Claude/ Gemini/ Llama or other model.	BRD-8, 9,10,11,12,13,14,15,16,26, 36
SDD-19	Data protection	GDPR/ CCPA and other applicable data protection.	BRD-54