LEO CDP Documentation

None

Table of contents

1. LEO CDP

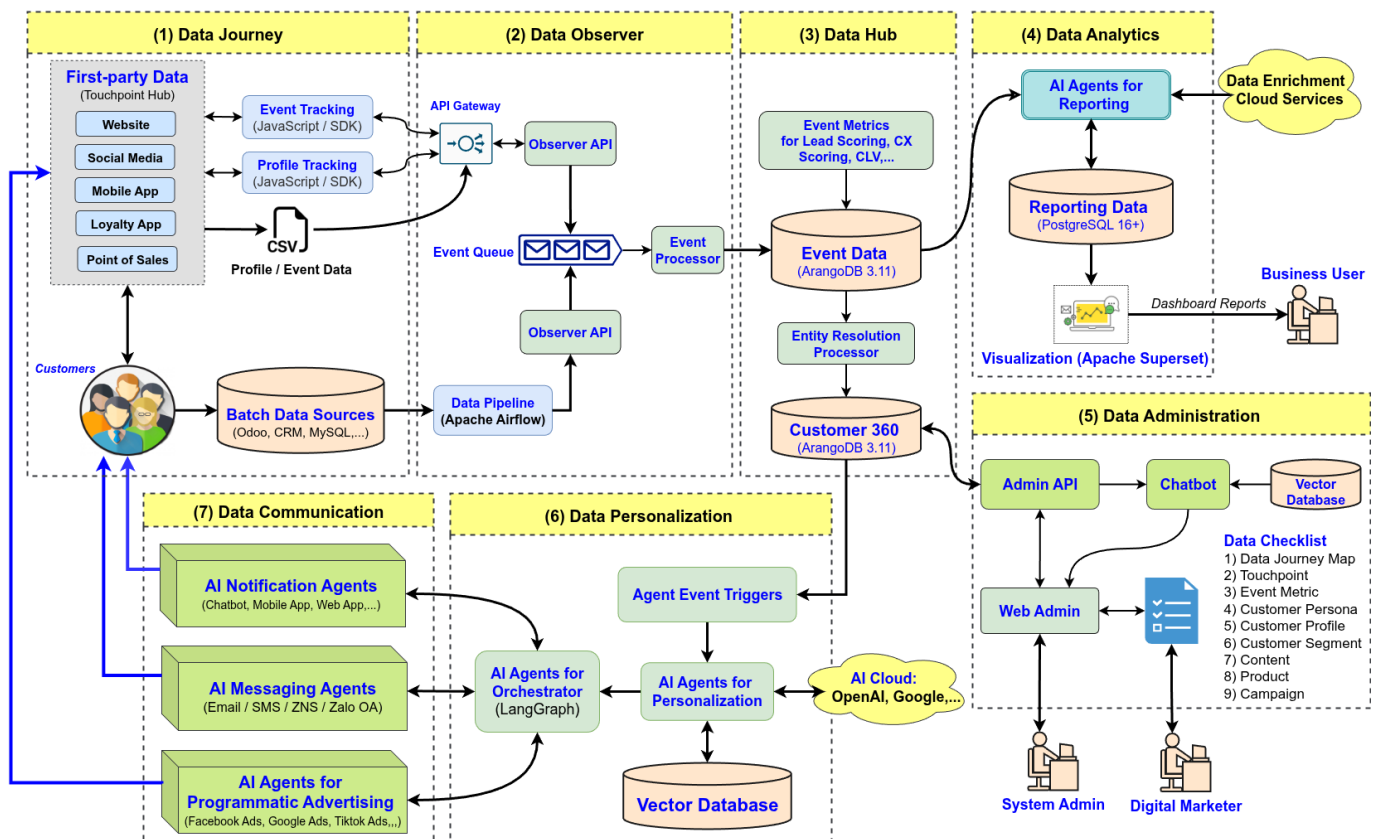## 1.0.1 An Open-Source, AI-First Customer Data Platform

LEO CDP is an **open-source, AI-first Customer Data Platform (CDP)** designed to help organizations build, own, and operate their **customer intelligence infrastructure** with full control over data, models, and execution.

Built with **machine learning, big data, and event-driven architecture at its core**, LEO CDP enables unified data collection, real-time analytics, intelligent segmentation, and AI-powered activation—while remaining **self-hosted, extensible, and privacy-respecting by design**.

LEO CDP is engineered for **developers, data engineers, data scientists, and modern marketing teams** who require flexibility, transparency, and long-term scalability beyond closed SaaS platforms.

AI-first Customer 360 Framework

Author: Nguyễn Tấn Triều (Thomas)
Copyright 2025

1.1 🚀 Vision & Philosophy

LEO CDP is grounded in a clear philosophical direction:

- **From Dataism → USPA → LEO CDP**: data as a first-class asset, intelligence as a system capability
- Democratize **AI-powered data platforms** for real digital transformation
- Promote **data sovereignty**, **on-premise intelligence**, and **open collaboration**
- Shift from rule-based marketing to **agentic, intent-driven customer engagement**

LEO CDP is not just a tool—it is an **architectural framework** for long-term customer intelligence.

1.2 🔥 Core Capabilities

1.2.1 1. Omnichannel Data Collection & Unification

Ingest customer data from web, mobile, CRM, POS, IoT, social platforms, and APIs. Normalize and unify events into a consistent, extensible data model.

1.2.2 2. Real-Time Customer 360

Construct a continuously updated customer profile combining behavioral, transactional, and contextual data across all touchpoints.

1.2.3 3. AI-Driven Segmentation & Scoring

Apply clustering, RFM analysis, CLV prediction, churn detection, and dynamic audience generation using machine learning models.

1.2.4 4. Behavioral Tracking & Journey Mapping

Track fine-grained user actions in real time and model customer journeys across channels and devices.

1.2.5 5. Predictive Analytics & Insights

Build and operate ML pipelines using Python, Jupyter, or Colab to generate actionable insights directly from CDP data.

1.2.6 6. Personalization & Activation (Agentic AI)

Leverage agentic AI to orchestrate personalized experiences across email, push notifications, SMS, and content delivery—driven by customer intent rather than static rules.

1.2.7 7. Event-Driven Architecture with ETL / ELT

Native support for event-based ingestion and transformation, with Apache Airflow integration for orchestration at scale.

1.2.8 8. API-First & Plugin-Oriented Design

Modular, extensible services with clean REST APIs—designed for integration, automation, and customization.

1.2.9 9. Data Governance & Privacy by Design

Built-in consent management, GDPR alignment, and on-premise deployment options to ensure full ownership and compliance.

1.2.10 10. DevOps-Ready by Default

Containerized deployment, observability with Prometheus and Grafana, and a scalable microservice architecture suitable for enterprise environments.

1.3 📈 Roadmap (2026+)

Feature	Status
Core CDP Platform (Profiles, Events, Segmentation)	✅ Complete
CDP SDKs (JavaScript, Python)	✅ Complete
SSO Integration with Keycloak	✅ Complete
AI Assistant for Audience Insights & Recommendations	🔄 In Progress
Knowledge Graph + Vector-Based Matching	🔄 In Progress
Agentic AI for Personalized Customer Experiences	🔄 In Progress
Webhooks & Event Bus Integration (Kafka)	🆕 Planned

1.4 🛠️ Technology Stack

- **Backend**: Java 11 (Amazon Corretto), Python 3.10
- **Databases**:
- ArangoDB 3.11.14 (Multi-model: Document, Graph, Search)
- **Data Pipeline**: Apache Airflow
- **Messaging & Caching**: Redis 8
- **Monitoring & Observability**: Prometheus 2, Grafana 8
- **Deployment**: Ubuntu 22 LTS, Docker, On-Premise or Cloud

LEO CDP is built for teams who believe that **customer intelligence should be owned, explainable, and evolvable**—not locked inside opaque platforms.

2. Architecture

3. LEO Activation

4. LEO CDP API

5. Tutorials

5.1 LEO CDP Get Started

5.2 LEO CDP - Advanced Features

6. LEO CDP - FAQ

7. Changelog

- 12/12 -