HubDock

OCMA Open Cloud MDM Alliance

PROJECT DEFINITION

HUB & DOCK Open Industry Standard for Master Data Management

V 1.0

Table of Contents

1 Vision

OCMA envisions a future where the complexities of Master Data Management (MDM) are eliminated through Plug-and-Play simplicity and community-driven innovation. Supported by HubDock and other founding members, OCMA leads the development of an open, scalable, and sustainable MDM ecosystem.

2 Short Description:

Open Cloud MDM, developed and governed by OCMA, is an open-source Master Data Management system designed to surpass traditional proprietary solutions like IBM and Informatica. HubDock, as a founding member, contributes technical concepts and resources to support this mission.

3 License:

GPLv3

4 Objectives

1. **Develop a Robust MDM**: Under the leadership of OCMA, create a scalable and modular MDM solution that serves as the backbone for business applications.
2. **Enable Plug-and-Play Integration:** Offer out-of-the-box compatibility with leading systems like SAP, Salesforce, and Microsoft Dynamics.
3. **Foster Community Innovation:** Leverage collective intelligence to co-create solutions tailored to real-world challenges.
4. **Establish an Open Industry Standard:** Establish an Open Industry Standard: Led by OCMA, define excellence in MDM through widespread adoption and recognition, supported by founding members like HubDock.
5. **Enhance Interoperability:** Provide native connectors and APIs for seamless integration into existing IT landscapes.

5 Tech Stack

The Open Cloud MDM platform leverages technologies such as Microsoft .NET, ASP.NET, and SQL Server. This technological foundation is designed and governed by OCMA, with contributions from partners like HubDock to ensure high standards and scalability.

- C# for business logic
- Microsoft SQL Server for data storage
- Utilizing ASP.NET for backend services and Angular for the frontend
- Maui for native desktop interface
- Docker and Kubernetes for containerised deployments

6 Documentation & Location

Publicly available documentation is available at documentation.hubdock.com.

Project documentation and deliverables are documented on GitHub. Access to team repositories is restricted to HubDock stakeholders only.

7 Key Features

Plug-and-Play Architecture
- Pre-configured connectors for popular platforms and systems
- Rapid deployment with minimal configuration requirements
- Automated discovery of data sources and targets

Modular and Scalable Design
- Horizontal and vertical scalability to meet evolving demands
- Modular components that simplify maintenance and upgrades

AI-Enhanced Data Management
- Machine learning algorithms for data cleansing, anomaly detection, and predictive analytics
- Real-time data synchronisation with built-in error recovery

Community-Driven Plugin Marketplace
- Dynamic marketplace for user-contributed plugins and extensions
- Incentives for community members to share high-quality contributions

Multi-Tenancy and Hybrid Deployment
- Support for cloud, on-premise, and hybrid environments
- Multi-tenant capabilities for SaaS providers

Advanced Security and Data Protection
- Robust encryption for data at rest and in transit
- Role-based access control (RBAC) and multi-factor authentication (MFA)

Workflow Automation and User-Centric Design
- Intuitive interfaces to reduce manual errors
- Automatable workflows to streamline data operations

Open Source Community Collaboration
- Extensive contributions and updates from an active open-source community
- Transparent development process with publicly available code and resources

Partnerships and Community Involvement

OCMA actively seeks strategic partnerships with academic institutions like Fraunhofer FIT and industry players to support the development of the Open Cloud MDM Core. Founding

members, including HubDock, facilitate these collaborations through technical and strategic contributions. These collaborations will:

- **Support MVP Development:** Facilitate funding and technical validation.
- **Accelerate Market Entry:** Enhance the platform's credibility and adoption.
- **Enable Co-Creation:** Foster joint innovation for long-term scalability.

8 Industry Standards

HubDock is dedicated to aligning with key industry standards to ensure consistency and quality across our platform. We strive to adhere to these standards as closely as possible, recognizing that in certain instances, adaptations may be necessary to accommodate the unique dynamics of our ecosystem and to resolve any potential contradictions between standards.

- **Integration of ISO 8000**
  We aim to ensure data quality and consistency across our platform by aligning with ISO 8000 standards for master data management. ISO 8000 is an international standard developed by the International Organization for Standardization.
- **Adoption of DAMA-DMBOK Best Practices**
  Our approach to data management is holistic, incorporating DAMA-DMBOK's comprehensive best practices to enhance data integrity and value. DAMA-DMBOK is provided by the Data Management Association International (DAMA International), and more information can be found on their website.
- **Alignment with GS1 for Global Consistency**
  We strive for global interoperability, particularly in supply chain management, by integrating GS1 standards for product identification and data synchronization. GS1 is a global organization, and more information about their standards can be found here.
- **Utilizing UNSPSC for Standardized Classification**
  To ensure consistent categorization of products and services, we aspire to adopt the UNSPSC taxonomy within our MDM solution. The United Nations Standard Products and Services Code (UNSPSC) is managed by GS1 US for the UN Development Programme.
- **FAIR Data Principles**
  These principles focus on making data Findable, Accessible, Interoperable, and Reusable, providing a framework for good data management and stewardship. More information on FAIR Data Principles can be found through the FORCE11 community group.
- **Schema.org**
  Schema.org is a collaborative, community activity with a mission to create, maintain, and promote schemas for structured data on the Internet, on web pages, in email messages, and beyond. Founded by Google, Microsoft, Yahoo and Yandex, Schema.org vocabularies are developed by an open community process, using the public-schemaorg@w3.org mailing list and through GitHub.
- **COBIT (Control Objectives for Information and Related Technologies)**
  COBIT is a framework for developing, implementing, monitoring, and improving IT

governance and management practices. It is provided by ISACA, an international professional association focused on IT governance. More information about COBIT can be found on the ISACA website.

9 General Requirements

9.1 Multi-Tenant Scalability

The **HubDock Open Cloud MDM** system, herein referred to as "The Hub," must be architected to support multi-tenant capabilities. This transcendent feature allows The Hub to serve not merely as an isolated instance but as an empowering, SaaS-enabling server within a cloud service environment. This seamless infrastructure should provide the backbone for SaaS using companies, facilitating their ability to host, manage, deploy, and integrate their applications in the cloud environment. This harmonized alignment ensures a community-driven, collaborative innovation while maximizing mutual benefits for all stakeholders in the ecosystem.

9.2 Hybrid MDM Capability

The system shall offer a harmonized, hybrid Master Data Management (MDM) architecture that seamlessly integrates with both on-premises and cloud-based data sources. This feature will empower organizations to achieve real-time data synchronization across diverse deployment models. The hybrid MDM must also support multi-tenancy configurations, enabling it to serve as the backbone for MDM SaaS providers.

9.3 Developer-Friendly API

A feature-rich, easy-to-navigate API that empowers developers to seamlessly integrate and extend the platform's capabilities.

9.4 AI-Enhanced Data Management

Leveraging artificial intelligence to enhance data quality, pattern recognition, and predictive analytics, transforming data into actionable insights.

9.5 Modular Architecture

A modular architecture that simplifies the upgrade and maintenance process, enabling an agile response to the rapidly evolving marketplace.

9.6 Security and Data Protection

The system's architecture must be meticulously designed and implemented to ensure the highest standards of security and data protection, safeguarding all user data and interactions within the ecosystem. This encompasses robust encryption, secure access controls, and continuous monitoring to pre-emptively identify and mitigate potential vulnerabilities, ensuring the integrity, confidentiality, and availability of data at all times.

9.7 Real-Time Data Synchronization and Error Recovery

Real-time data sync across all nodes with built-in error recovery mechanisms, ensuring data consistency and system resilience.

9.8 Emphasis on Reusability and Maintainability

A robust set of core components and libraries aimed at reducing technical debt and making the system future-proof.

9.9 User-Centric Design and Workflow Automation

An intuitive user interface complemented by workflow automation capabilities, to enhance productivity and minimize manual errors.

9.10 Community-Driven Plugin Marketplace

A dynamic, community-driven marketplace encouraging collaborative innovations and extending the system's functionalities through high-quality plugins.

9.11 Enhanced Scalability

A system architecture designed for horizontal scaling, ready to meet the growing demands of an expanding user base.

9.12 Simplified Job Scheduling and Enhanced Error Handling

A sophisticated job scheduler with robust error-handling capabilities, streamlining routine tasks and enhancing system reliability.

9.13 Rewarding Collaborative Innovations

An incentivization mechanism to reward community contributions, thereby promoting collaborative innovations within the ecosystem.

10 Scope

10.1 Hub Interfaces: Administration Interfaces

Designed to offer intricate controls and advanced customization settings, enabling ecosystem partners to tailor their administrative interface to specific needs.

Scope: To design and implement a user-friendly yet robust administration interface.

Details: This will allow administrators to manage data schemas, user permissions, and system configurations effortlessly, setting the stage for a community-driven administration experience.

10.2 MDM Workbench

This is the cockpit for data engineers and administrators. By leveraging the strengths of a hybrid MDM, the workbench will empower users to manage, monitor, and manipulate data in real-time.

Scope: To develop a comprehensive workbench for data management tasks.

Details: Equipped with tools for data profiling, quality checks, and workflow automation, the workbench will serve as the central hub for data administrators and engineers.

10.3 Request Framework

Built on a robust, yet flexible architecture, this framework orchestrates all data requests, offering both synchronous and asynchronous capabilities. As a cornerstone of our hybrid approach, it manages transactions for operational data and analytical insights.

Scope: To create a flexible, efficient request framework for handling data requests.

Details: This will manage both read and write operations, synchronous and asynchronous data transactions, offering a seamless experience across the ecosystem.

10.4 Batch Processing Framework

A refined batch processing mechanism that automates large-scale data operations without compromising performance, thereby balancing the needs of real-time operational data with large-scale analytical processes.

Scope: To enable high-volume data import, export, transformation, and validation tasks in batch mode.

Details: This will manage both read and write operations, synchronous and asynchronous data transactions, offering a seamless experience across the ecosystem.

10.5 DOCK - The Hub's Integration Framework

DOCK acts as the dynamic nerve center that interconnects various data streams, applications, and services. Its modular design supports a wide range of protocols and data formats, making it a key element in the hybrid MDM approach.

Scope: To establish a versatile integration framework called DOCK, an universal connector architecture that allows seamless integration with external systems through RESTful APIs, Webhooks, and SDKs.

Details: This will support various data exchange protocols and formats, providing a modular, plug-and-play architecture for data integration and transformation.

10.6 Core Components

These are the fundamental elements upon which the entire ecosystem is built. From data ingestion to normalization and governance, the core components are designed to be both robust and scalable.

Scope: The building blocks of the system, including the master data model, data storage services, and business logic handlers.

Details: This includes components for data ingestion, normalization, governance, and dissemination. The Core Components are foundational and must be reliable and scalable.

10.7 Extension Framework

A sophisticated layer allowing for data and rule extensions, this is where the ecosystem thrives on collaborative innovations. It empowers users to build upon the existing core components, allowing the MDM system to adapt and evolve over time.

Scope: To develop an extension framework for data and rule customization.

Details: This will allow ecosystem partners to extend and enhance functionalities, promoting collaborative innovations while keeping the core intact.

10.8 Common Components

These include the non-functional elements like security, logging, and monitoring. They bring mutual benefits by ensuring that the MDM system is not just powerful, but also secure and easy to manage.

Scope: To implement standard components that handle non-functional requirements, reusable modules for logging, authentication, and other cross-cutting concerns.

Details: Features like security, logging, and monitoring will be uniform across the ecosystem, providing mutual benefits in manageability and security.

11 Development Procedure

In line with our philosophy of collaborative innovation and agile responsiveness, we adopt an iterative and incremental approach to development. This methodology enables us to adapt dynamically to the needs of our ecosystem while ensuring quality and reliability.

11.1 Initial Development: Prototype (April 2025 – September 2025)

Assemble the core team, including a CTO and Lead Developer.

Focus on delivering a prototype showcasing essential functionalities, such as:

- Plug-and-Play integration for seamless system compatibility,
- Core data models and validation frameworks,
- Administrative interfaces for configuration and basic system management.

This prototype will validate the architectural choices and serve as a proof of concept for potential partners and stakeholders. It sets the foundation for the Minimum Viable Product (MVP) by establishing a scalable architecture and identifying initial technical challenges.

11.2 First Iteration: Minimum Viable Product and Community Integration (MVP, October 2025 – March 2026)

Engage with strategic industry partners to ensure broad adoption and real-world validation. Develop the Minimum Viable Product (MVP), focusing on:

- Core components like data ingestion, secure storage, and an operational API framework.
- Launch the Plugin Marketplace to enable community-driven innovation.
- Refine the user interface for improved usability.
- Integrate AI-powered data management tools for advanced data cleansing and anomaly detection.

Feedback from early adopters will be systematically gathered to prioritise future. The MVP phase transitions into the next iteration by incorporating real-world insights, which guide the optimisation and scalability efforts.

11.3 Second Iteration: Scalability and Industry Adoption (April 2026 – December 2026)

Optimise the platform for larger-scale deployments, focusing on:

- Enhanced system reliability, including error recovery and advanced logging mechanisms.
- Intuitive and customisable user interfaces tailored for various user roles.
- Introducing industry-specific extensions through the Plugin Marketplace.

Scope: This iteration ensures the platform meets the demands of large-scale enterprise environments, paving the way for industry-wide adoption. It builds on the insights from the MVP phase, refining both technical stability and user experience to support broader market requirements.

11.4 Third Iteration: Advanced Analytics and Monetisation (2027)

Launch monetisation models, including:

- Subscription-based MDM as a Service.
- Premium licensing options for enterprises requiring on-premise solutions.

Introduce advanced data analytics by integrating Microsoft Power BI features:

- Sophisticated reporting and real-time insights for strategic decision-making.
- Advanced visualisation capabilities to support data-driven business processes.

Scope: This iteration solidifies HubDock as a comprehensive, data-driven decision-making platform, setting a new standard in Master Data Management. By leveraging the stability and scalability achieved in previous phases, this stage introduces monetisation and advanced features to maximise value for users and stakeholders.

The proposed timeline for the iterative development of "The Hub" serves as a guiding framework and is subject to change. The pace of our progress is influenced by a constellation of factors, chief among them being the level of community involvement, availability of development resources, and the financial support garnered from sponsors and early adopters.

By aligning with HubDock, you're not just investing in a product but becoming part of a visionary endeavor to shape the future of Master Data Management. While we are steadfast in our commitment to adhere to these timelines as closely as possible, we recognize the dynamic nature of software development within a community-driven ecosystem. Adjustments to this timeline may be necessitated to ensure that the platform we co-create is not just robust, but also intricately tailored to real-world requirements and challenges.

Your engagement and contributions are invaluable to us. The timelines will be updated to reflect any major pivots or milestone achievements, and we remain committed to transparent communication every step of the way.

In alignment with HubDock's vision for iterative and incremental development, the following milestones articulate the phased evolution of "The Hub." These milestones serve as checkpoints that not only assess our progress but also reiterate our commitment to delivering an empowering, community-driven MDM system.

12 Budget and Funding

The MVP development will be financed through OCMA's funding streams, including public grants, membership fees, and sponsorships. Founding members like HubDock contribute financial resources and early-stage investments to accelerate progress.

13 HR Management:

1. **Core Team**: Comprising of maintainer, architects, and lead developers.
2. **Community Contributors**: Vetted developers and experts contributing to various modules.
3. **Academic Partnerships:** Leverage expertise from institutions like Fraunhofer FIT
4. **Quality Assurance**: A separate team focused on maintaining the high quality of the project.

14 Risk Assessment:

1. **Code Quality**: Mitigated by strict review processes.
2. **Community Engagement**: Ensured through regular updates and transparent communication.
3. **Resource Constraints**: Addressed by scaling the team based on community contributions and potential partnerships.

15 Constraints:

1. **Time**: Achieving milestones within the stipulated timelines.
2. **Quality**: Ensuring the end product meets or exceeds current industry standards.

16 Success Criteria:

1. **Adoption Rates**: High adoption by enterprises, developers, and HubDock ecosystem partners.
2. **Community Contributions**: Regular, high-quality contributions from the community.
3. **Performance Metrics**: Outperforming current industry leaders in key metrics.
4. **Industry Recognition:** Establish HubDock as the de facto standard for open-source MDM.

17 Communication Plan:

1. **Monthly Updates**: Via GitHub and HubDock community channels.
2. **Quarterly Reviews**: In-depth review sessions open to all stakeholders.

18 APPENDIX

19 HubDock Code of Conduct

At HubDock, we are committed to fostering a supportive and collaborative ecosystem. Our Code of Conduct reflects our values and sets the standard for how we interact and work together.

Personal Criteria:

- **Active Participation**: Engage consistently and contribute to our shared goals.
- **Eagerness to Learn & Excel**: Always be curious and strive to improve.
- **Team Spirit**: Work together and celebrate our collective achievements.
- **Supportive Nature**: Lend a hand to others on their journey to success.
- **Transparency, Honesty, & Reliability**: Build trust through open and reliable actions.
- **Commitment**: Show dedication to our mission and values.
- **Courtesy**: Maintain respect in all interactions.

Professional Criteria:

General:

- Bring at least two years of experience in MDM, CRM, or ERP application design or development, contributing to our pool of knowledge.

For Developers:

- Have a good grasp of C# and ASP.NET, creating robust solutions.
- Be familiar with "MAUI" and "ANGULAR" frameworks, adapting to different challenges.

Guiding Principles:

- **Respect**: We value each other and acknowledge everyone's contributions.
- **Integrity**: We operate transparently and hold ourselves accountable.
- **Inclusivity**: We embrace diversity and ensure everyone feels included.
- **Collaboration**: We believe in the power of working together and sharing ideas.