

PROJECT DEFINITION

HUB & DOCK

Open Industry Standard for Master Data Management

V 1.0





Table of Contents

- 1 Vision 5
- 2 Short Description:..... 5
- 3 License:..... 5
- 4 Objectives 5
- 5 Tech Stack..... 5
- 6 Documentation & Location 5
- 7 Industry Standards 6
- 8 General Requirements 7
 - 8.1 Multi-Tenant Scalability..... 7
 - 8.2 Hybrid MDM Capability..... 7
 - 8.3 Developer-Friendly API 7
 - 8.4 AI-Enhanced Data Management 7
 - 8.5 Modular Architecture 7
 - 8.6 Security and Data Protection..... 7
 - 8.7 Real-Time Data Synchronization and Error Recovery..... 7
 - 8.8 Emphasis on Reusability and Maintainability 8
 - 8.9 User-Centric Design and Workflow Automation 8
 - 8.10 Community-Driven Plugin Marketplace..... 8
 - 8.11 Enhanced Scalability 8
 - 8.12 Simplified Job Scheduling and Enhanced Error Handling..... 8
 - 8.13 Rewarding Collaborative Innovations 8
- 9 Scope..... 8
 - 9.1 Hub Interfaces: Administration Interfaces 8
 - 9.2 MDM Workbench..... 8
 - 9.3 Request Framework..... 9
 - 9.4 Batch Processing Framework..... 9
 - 9.5 DOCK - The Hub's Integration Framework 9
 - 9.6 Core Components 9
 - 9.7 Extension Framework..... 10
 - 9.8 Common Components 10
- 10 Development Procedure 10
 - 10.1 First Iteration: Minimum Viable Product (MVP) 10
 - 10.2 Second Iteration: Enhanced User Experience and Reliability 10
 - 10.3 Third Iteration: Scalability and Extensibility..... 11
 - 10.4 Fourth Iteration: Data Fabric Features Built on Microsoft Power BI 11
- 11 Timeline and Milestones 11



- 11.1 Initial Pilot Release..... 11
- 11.2 Enhancements and Community Feedback..... 12
- 11.3 Hybrid MDM and API Development..... 12
- 11.4 Data Fabric Integration..... 12
- 11.5 First Annual Review..... 12
- 12 HR Management:..... 12
- 13 Risk Assessment: 13
- 14 Budget: 13
- 15 Constraints: 13
- 16 Success Criteria:..... 13
- 17 Communication Plan:..... 13
- 18 APPENDIX 14
- 19 HubDock Code of Conduct 15

1 Vision

We envision a future where the complexities of master data management are dissolved through community-driven innovation. HubDock aims to transcend existing paradigms, creating an open-source MDM solution that serves as the bedrock for a harmonized data landscape across industries.

2 Short Description:

Creating an open-source Master Data Management system, the **HubDock Open Cloud MDM**, designed to transcend existing solutions like IBM and Informatica, thereby becoming an open industry standard. This platform is not merely a tool, but a powerful catalyst for innovation, fostering seamless integration and harmonization across diverse sectors. By leveraging the 'Hub', organizations can unlock the full potential of their data, ensuring accuracy, consistency, and accessibility.

Learn more at [HubDock - Empowering Businesses](#)

3 License:

GPLv3

4 Objectives

1. **Develop a Robust MDM:** Craft an MDM that will act as the backbone for business applications, ensuring seamless integration and data consistency.
2. **Community Involvement:** Leverage collective intelligence for the co-creation of the MDM, ensuring solutions are tailored to real-world needs.
3. **Open Industry Standard:** Aim to be a benchmark in open-source MDM, defining clear excellence and innovation criteria. Success will be marked by widespread adoption and recognition as a leading resource in this domain.

5 Tech Stack

- 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

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 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](#).

8 General Requirements

8.1 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.

8.2 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.

8.3 Developer-Friendly API

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

8.4 AI-Enhanced Data Management

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

8.5 Modular Architecture

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

8.6 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.

8.7 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.

8.8 Emphasis on Reusability and Maintainability

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

8.9 User-Centric Design and Workflow Automation

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

8.10 Community-Driven Plugin Marketplace

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

8.11 Enhanced Scalability

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

8.12 Simplified Job Scheduling and Enhanced Error Handling

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

8.13 Rewarding Collaborative Innovations

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

9 Scope

9.1 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.

9.2 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.

9.3 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.

9.4 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.

9.5 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.

9.6 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.

9.7 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.

9.8 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.

10 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 be dynamic in meeting the needs of our ecosystem while assuring quality and reliability.

10.1 First Iteration: Minimum Viable Product (MVP)

The primary focus of our first iteration will be to deliver a Minimum Viable Product (MVP) that encapsulates the essential functionalities of "The Hub." This iteration aims to create a functioning system that meets the most critical needs, acting as a launchpad for subsequent refinements and feature additions.

Scope: To build the backbone of "The Hub," focusing on core components like data ingestion, storage, and the basic administrative interface.

10.2 Second Iteration: Enhanced User Experience and Reliability

As we move to the second iteration, our focus will shift to enhancing the user experience and system reliability. This includes developing a more intuitive user interface, more robust error-handling capabilities, and introducing initial elements of our community-driven plugin marketplace.

Scope: To refine and extend the capabilities of the MVP, based on real-world feedback and performance metrics.

10.3 Third Iteration: Scalability and Extensibility

The third iteration will pivot toward enhancing the system's scalability and extensibility. By aligning with HubDock's vision of a community-driven ecosystem, this iteration aims to add advanced features such as AI-enhanced data management and batch processing frameworks, setting the stage for collaborative innovations.

Scope: To optimize "The Hub" for scalability, prepare for larger user-base adoption, and integrate community contributions into the core system.

10.4 Fourth Iteration: Data Fabric Features Built on Microsoft Power BI

In this advanced phase, we turn our attention to data analytics and business intelligence by integrating Data Fabric features built on Microsoft Power BI. This serves as an exemplar of HubDock's commitment to seamless integration and real-time insights. The addition of these features amplifies the MDM's capabilities, transforming it into a holistic, data-driven decision-making tool.

Scope: To integrate Microsoft Power BI capabilities into "The Hub," enabling more intricate data analytics, reporting, and visualization functionalities. This will set the stage for providing real-time, actionable insights for businesses who are part of the HubDock ecosystem.

11 Timeline and Milestones

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.

11.1 Initial Pilot Release

- **Timeline:** 01/2024 .. 03/2024
- **Scope:** Deliver the core architecture, focusing on a minimally viable product that integrates essential MDM functionalities.

- **Deliverables:** Initial codebase, documentation, and preliminary community involvement through beta testing.

11.2 Enhancements and Community Feedback

- **Timeline:** 04/2024 .. 05/2024
- **Scope:** Refine features and functionality based on community feedback, focusing on usability and extendability.
- **Deliverables:** Feature updates, more comprehensive documentation, and initiation of a community-driven plugin marketplace.

11.3 Hybrid MDM and API Development

- **Timeline:** 06/2024 .. 12/2024
- **Scope:** Introduce the hybrid MDM capabilities and a more developer-friendly API to foster ecosystem growth.
- **Deliverables:** A robust API, documentation for developers, and the introduction of hybrid MDM functionalities.

11.4 Data Fabric Integration

- **Timeline:** 01/2025 .. 06/2025
- **Scope:** Build on Microsoft Power BI to integrate Data Fabric features, thereby enhancing the system's analytical capabilities.
- **Deliverables:** Power BI integration, advanced analytics tools, and updated documentation.

11.5 First Annual Review

- **Timeline:** End of 2024
- **Scope:** Evaluate the progress of the project, with an emphasis on user adoption rates, community contributions, and system performance.
- **Deliverables:** A comprehensive annual report, highlighting achievements, lessons learned, and future roadmaps.

12 HR Management:

1. **Core Team:** Comprising of maintainer, architects, and lead developers.
2. **Community Contributors:** Vetted developers and experts contributing to various modules.

3. **Quality Assurance:** A separate team focused on maintaining the high quality of the project.

13 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.

14 Budget:

To be financed by HubDock and potentially through community crowdfunding, grants, or 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.

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.