Semarchy

Company & Product Overview

Complimentary eBook

2021

Contents

Company

Overview

Semarchy is the Intelligent Data Hub™ company. Its xDM platform solves for enterprise Master Data Management (MDM), Application Data Management (ADM), and Collaborative Data Governance challenges at some of the most well-known brands across the US and Europe. This agile platform leverages smart algorithms and Google's Material Design UI standards to empower any data steward, business user, or even third party, to intuitively own and manage relevant aspects of their business-critical data.

Our xDM data hub combines comprehensive features covering automated data discovery, integration, stewardship, data quality and enrichment, workflows, and core MDM capabilities, alongside in-depth metrics and dashboards all in a single product. *x*DM from Semarchy was built fully in-house, on a carefully designed architecture and technology stack to allow rapid implementation via an agile and iterative approach that delivers business value almost immediately yet scales to meet enterprise complexity.

Semarchy was founded after the consolidation in the MDM market circa 2010. Our founders felt that there was a major gap—picking up where that technology leaves off, with data integration, governance, quality, and workflows as part of a single platform solution. At the same time, there was a business case gap--around time to value, agility, and meeting business requirements with data management applications and projects.

More about the Semarchy Team can be found on
https://www.semarchy.com/management-team.

Intelligent Data Hub™ - Market Positioning

Semarchy has been updating its overall market message to keep pace with the evolution of the Master Data Management space. The software platform itself has expanded capabilities consistently, in line with the initial objectives of the business: to provide organizations with a single platform for multi-vector MDM, with in-built information stewardship/governance, data quality and enrichment, workflows, dashboards, hierarchy management, and reference data management in parallel.

In years past, the company explained its products as "Evolutionary MDM," then "Intelligent MDM", as smart algorithms and semantic learning have been incorporated into the platform. The company has trademarked both terms. Semarchy refers to xDM as an "Intelligent Data Hub™."

As we continue to understand why organizations invest in Semarchy, and consistently renew and augment their investment, our understanding of the value provided has become more focused.

Semarchy explains the platform broadly as Smart, Agile, and Measurable:

Smart

Algorithms empower stewardship and advance matching, survivorship, curation, and classification. A native, semantic model facilitates transparent lineage, auditability, and governance. Like Gartner, Semarchy believes that MDM is about delivering enterprise-centric semantic consistency for any kind of data, and in a single platform.

Agile

Agile methodologies applied to MDM are materialized via an intuitive, integrated design, development and user interface. Adaptive modeling enables business flexibility. Semarchy has embraced the agile software development approach along with the Material Design Language popularized by Google and employs it for MDM to drive time-to-value.

Measurable

Like Gartner, Semarchy sees MDM as an initiative, not a technology. Following the seven building blocks for a successful MDM Plan, xDM implementations start with the business value before the technology. For this reason, xDM has been engineered so that it solves business problems with minimal disruption, but also evolves with the needs of the organization to optimize adoption and show measurable enterprise ROI almost immediately and increasingly so over time.

Intelligent Data Hub strategy in practice

The application of Agile methodologies forces collaboration between all key stakeholders to define and deliver solutions with value very rapidly. Our Proof-of-Value (PoV) process dovetails into a first iteration solution, which in turn grows via multiple fast iterations into a full working data hub quickly and efficiently. It's not unusual to hear client stakeholders comment that a PoV application "looks like a production solution."

Business Operations

Semarchy is a privately held company, founded in 2011. In 2016, we moved our strategic headquarters to San Francisco, California. We have active offices in Lyon, France (legal headquarters) and Maidenhead, UK with a further network of remote-work employees to extend our reach.

Semarchy operates via direct and partner channels in North America as well as Northern and Western Europe. The company is expanding via partner channels in South America, Russia, and Asia.

More information on the company mission, vision, and leadership can be found on our website.

Intelligent Data Hub™

Overview

The Intelligent Data Hub™ brings together the critical information that lives across applications such that it can be governed, mastered and managed in a centrally understood, non-disruptive way.
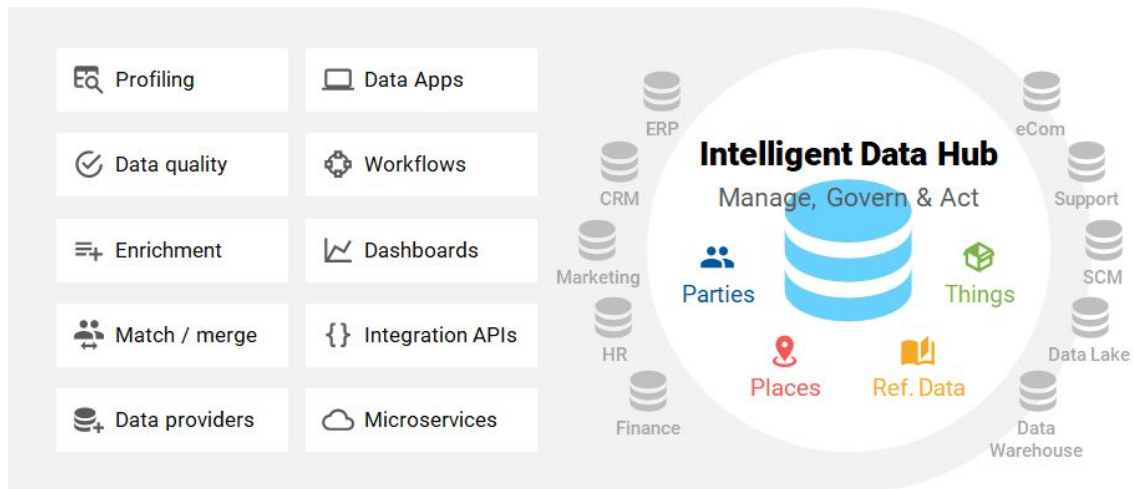
Intelligent Data Hub™ Architecture

The Intelligent Data Hub™ can be applied to several types of applications to solve critical business challenges to deliver value across the data management chain. At Semarchy, we believe that a data hub should provide the enterprise services to consistently manage the enterprises' core data about **Parties** (people, citizens, organizations, customers, suppliers, business partners, legal entities, employees, business units, etc.), **Places** (locations, addresses, facilities, stores, shops, warehouses, geographies, etc.) and **Things** (products, items, packages, equities, securities, assets, accounts, hierarchies, devices, etc.).

With the growing interactions and observations collected by enterprises through various data acquisition points, it is becoming critical to collaboratively govern these core data elements in an agile and measurable environment. Setting up this collaborative governance within an Intelligent Data Hub drives accurate interpretation of all the relationships between these core data elements.
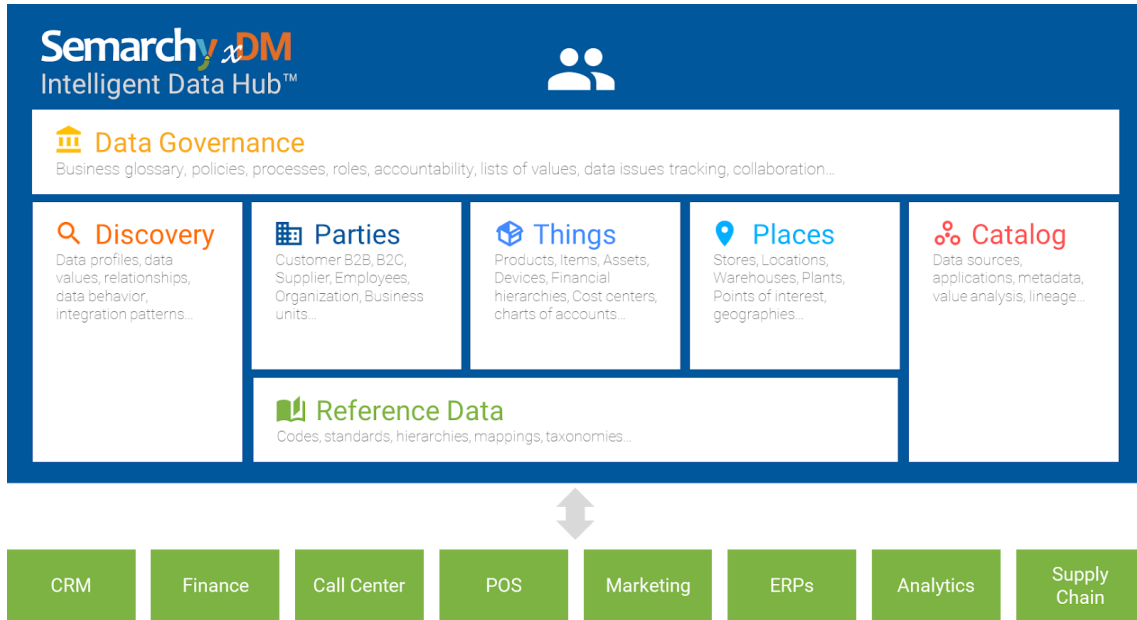
The Intelligent Data Hub™ Platform

With the Intelligent Data Hub™, operational analytics, predictive analytics, and traditional analytics share the same master data, reference data, business glossary and data catalogs to transform the interactions and observations into measurable business value. Similarly, operational business applications leverage the Intelligent Data Hub in their core processes to guarantee semantic consistency across the enterprise, thus reducing costs and risks and increasing revenue opportunities.
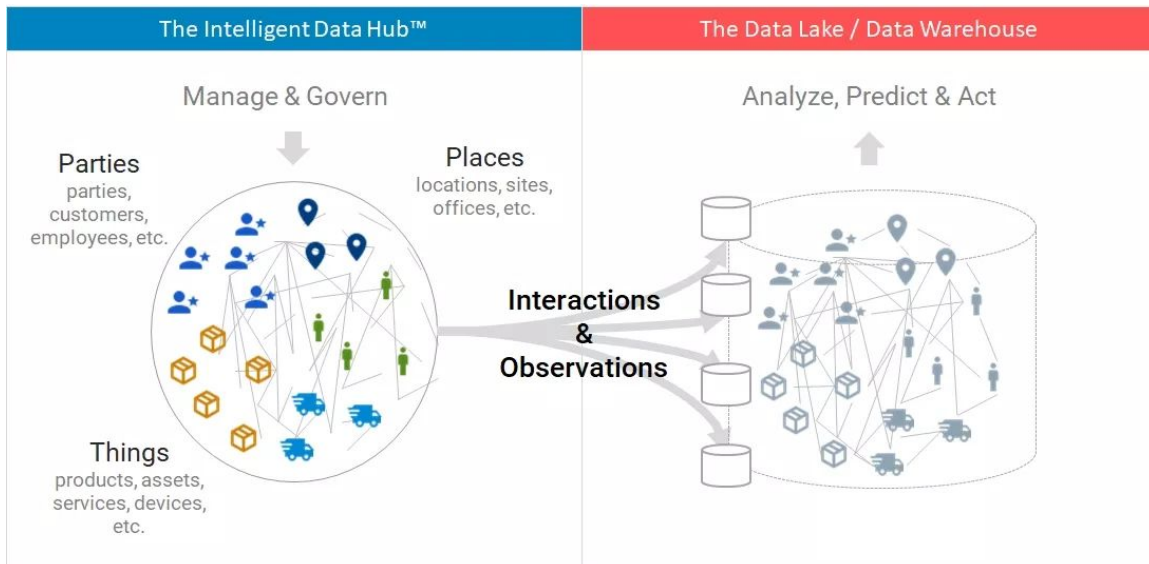
The Intelligent Data Hub™ for analytics.

Semarchy clients benefit from our platform to design and deploy critical applications orchestrating business processes across multiple lines of businesses, all sharing a single consistent and governed data hub.

Discover how our clients use the Intelligent Data Hub from Semarchy on https://www.semarchy.com/clients.

Intelligent Data Hub for Parties

Customer B2B, Legal Entities, Business Partners, and Suppliers

The core data model for such a domain is usually centered on the **Legal Entity** (company) information and is augmented with information elements such as addresses, sites, hierarchies, market segments, contacts, financials, history, channels, risk metrics, markets, preferences, contracts, sales, services levels, ratings, loyalty, legal structure, compliance and other interactions and observations.

Critical business processes that can partially or fully be managed in the hub include new company creation, match & merge, deduplication, basic data quality, data survivorship and consolidation, external data enrichment, company data lifecycle, legal hierarchy management, advanced data quality / standardization, counter-party risk assessment (credit risk, legal / IP risk, financial risk), GDPR, HIPAA, marketing segmentation, surveys, retention, loyalty, and predictive analytics processes.

Implementation styles vary depending on the business requirements. They are detailed in the Multiple implementation style support / Multiple domain support.

Customer B2C, Citizens, and Households

The core data model for such a domain is usually centered on the Physical Person (consumer) information and is augmented with information elements such as addresses, channels, contact points, household, financials, credit risk, segmentation, credit history, legal data, preferences, social behavior, sales, contracts, and compliance.

Critical business processes that can be partially or fully managed in the hub include new person creation, match & merge, deduplication, basic data quality, data survivorship and consolidation, self-service preferences, person data lifecycle, external data enrichment, household management, advanced data quality / standardization, GDPR, CCPA, privacy, opt-in/out processes, credit risk, legal risk, financial assessment, marketing segmentation, surveys, retention, loyalty, predictive analytics, FoF circles, recommendations, multi-channel strategy.

Implementation styles vary depending on the business requirements. They are detailed in the [Multiple implementation style support / Multiple domain support](#).

Employees, HR, Business Units

The core data model for such a domain is usually centered on the **Employee** (whether internal or contractor) and is augmented with information elements such as addresses, reporting hierarchies, cost centers, business units, legal IDs, employment history, insurance, benefits, authentication, contracts, talents and skills, performance, salaries and bonuses, taxes, travel and expenses, recommendations, procurement, and social activity.

Critical business processes that can partially or fully be managed in the hub include new employee creation, match & merge, deduplication, basic data quality, data survivorship and consolidation, self-service enrichment, hiring approvals, multi-application provisioning, off-boarding, job mapping, re-org, reporting hierarchies, costs optimization, authentication, business continuity, taxes, environment, health, legal, privacy, screening assessments, GDPR, cost centers assignment, performance assessments, talent management, HR costs optimization, policies enforcement, referral programs

Implementation styles vary depending on business requirements. They are detailed in the [Multiple implementation style support / Multiple domain support](#).

Intelligent Data Hub for Things

Product Data (buy/sell-side)

The core data model for such a domain is usually centered on the **Product** (or Part/SKU) and is augmented with information elements such as extensible attributes, families and taxonomies, bill of materials, rich content (media, images, videos, pdfs, etc.), variants (sizes, colors, etc.), markets, geographies, pricing, lifecycle, branding and IP, legal data, compliance data, marketing campaigns, costs of sales, social metrics, inventory, supply chain information, sales, recommendations and call center statistics.

Critical business processes that can partially or fully be managed in the hub include product creation, deduplication, consolidation, collaborative information completeness, partner enrichment, classification, media enrichments, configuration, BOMs (MBOM, SBOM, EBOM), lifecycle management, hierarchies, advanced DQ, pricing, transfer prices, CFR 11, IDMP, MiFID, BCBS239, licensing, IP, environment, health, formulation, performance monitoring, profitability, quality insurance, satisfaction / loyalty surveys, recommendations, social campaigns, planning…

Implementation styles vary depending on business requirements. They are detailed in the [Multiple implementation style support / Multiple domain support](#).

Financial Instruments, Services, Hierarchies, and Securities

The core data model for such a domain is usually centered on the **Financial Instrument** and is augmented with information elements such as securities, tickers, policies, contracts, coverages, risks, costs, rating factors, hierarchies, pricing, etc.

Critical business processes that can partially or fully be managed in the hub include instrument creation, deduplication, consolidation, collaborative information completeness, pricing model enrichment, classification, configuration, lifecycle management, hierarchies, advanced DQ, legal compliance, performance monitoring, profitability, trading transparency, planning…

Implementation styles vary depending on business requirements. They are detailed in the [Multiple implementation style support / Multiple domain support](#).

Assets

The core data model for such a domain is usually centered on the **Asset** and is augmented with information elements such as location, cost center, amortization, tagging, insurance, debt, credits, risks, etc.

Critical business processes that can partially or fully be managed in the hub include asset creation, deduplication, consolidation, classification, configuration, lifecycle and timeline management, hierarchies, advanced DQ, financial compliance, legal compliance, amortizations, maintenance, planning, etc.

Implementation styles vary depending on the business requirements. They are detailed in the [Multiple implementation style support / Multiple domain support](#).

Intelligent Data Hub for Places

The core data model for such a domain is usually centered on the **Location** and is augmented with information elements such as geography, point of interest, administrative areas, public facilities, terrain, elevation, schedules, maintenance, rich content (pictures, videos, .pdf), contacts, facilities costs, business continuity, insurance, sales performance, taxes, legal data, compliance, quality insurance, supply chain, web traffic, planning, and social activity.

Critical business processes that can partially or fully be managed in the hub include location creation, deduplication, consolidation, information completeness, location public directory, marketing & media enrichments, internal technical support, stores/plant/warehouse readiness, business continuity, quality insurance, taxes assessment, legal risks, leases, financial assessment, privacy, GDPR, environment compliance, health compliance, location advertising, external exposure, business performance optimization & predictions, sales impact analysis, brand awareness…

Implementation styles vary depending on the business requirements. They are detailed in the [Multiple implementation style support / Multiple domain support](#).

Intelligent Data Hub for Reference Data

Reference Data Management use cases are part of the Intelligent Data Hub implementations. Typical implementations use the data hub for managing standards and codes. These include enterprise local standards, lists of values, history, industry standards, taxonomies, local mapping (application-to-application, application-to-hub, industry-standard-to-local-standard), version control, regulatory compliance, impact analysis and lineage, code values propagation and lifecycle.

Processes and touch points in the applications ecosystem include taxonomy governance, code-to-code translation, global versus local versus industry definitions, industry standard onboarding, outbound reporting, auditing, internal propagation, compliance enforcement, external exposure, business performance optimization & predictions, sales impact analysis, etc.

Most of our clients use xDM to manage their reference data, either in the context of a traditional domain (local reference data) or as a standalone central initiative often linked to the data governance processes.

Intelligent Data Hub for Data Governance, Data Discovery, and Data Catalog

xDM, as an Intelligent Data Hub includes pre-built applications for data governance, discovery, and data cataloging.

xDM includes a **Data Discovery** module to profile the data of any external application as well as the data that lives in the data hub to provide immediate KPIs and metrics on the data quality. This module is included as an extension to the Data Governance application and fully integrated with the other components.

The **Data Governance** application is customizable and includes a starter pack with

- an enterprise **glossary** to manage business terms, relationships, policies, processes, tasks, KPIs, domains
- a data **dictionary** to manage all physical data sets with their structural metadata and sample data for discovery
- **Users, roles, authorizations**, access contexts, risks, business stakeholders, and RACI matrices
- Data issues escalation workflows
- Advanced **data discovery and profiling**.

Refer to Information Governance and Policy Setting for additional details.

xDM – Packaged Enterprise Software Product

Semarchy has one main product offering, **Semarchy xDM**. Previous versions in the same code line were referred to as *Convergence for MDM*. Since version 4, the platform was renamed *xDM*. The current release version is **Semarchy xDM v5.2**.

Semarchy xDM platform at a glance

At a high level, xDM is the core platform that enterprises use to instantiate the Intelligent Data Hub™. It includes most of the features to **Discover**, **Integrate**, **Manage**, **Govern** the data and **Measure** the outcomes
through KPIs.

Discover
Connect to sources, profile data, discover critical assets, build data catalogs

Integrate
Real-time API integration, Batch loading APIs, External data plug-ins, Microservices

Manage
Apps for data champions & business users, built-in data quality, match/merge & more…

Govern
Business glossary, lineage, policy definition & enforcement, rules & processes

Measure
Dashboards & metrics based on any data, define ad-hoc KPIs and take actions

Key features recently introduced in Semarchy xDM include:

- **Intuitive Matching Algorithms**, including "Transitive Scoring" to achieve more accurate scores for the groups without modifying existing match rules, and Multi-Iteration Grouping to obtain better golden clusters, dramatically reducing the time spent in manual match and merge operations.
- **REST Clients** to support a zero-code approach to leverage third-party services for automated data enrichment, standardization and quality validation.
- **xDM Discovery**, a component for data architects and data champions to profile data sources to infer the structure and rules of the data hub model and to provide on-going metrics of the data quality with drill-down to the source data / root cause.
- **xDM Dashboards**, a component for data architects and data champions to simply design Dashboard Applications and provide business users with dynamic queries and charts based on data managed in the data hubs and any other data store. These charts and dashboards are embeddable into xDM data management applications or enterprise portals.
- **Comprehensive REST API** for data queries, integration and interactions, as well as for management and DevOps operations.
- **Cloud Databases and Multi-Database** certification for repositories and data hubs. xDM now supports Oracle database, PostgreSQL and Microsoft SQL on-premises or in most common cloud platforms (Amazon Web Services, Microsoft Azure and Google Cloud Platform)

Key features in xDM can be found here: semarchy.com/xdm/. Extensive documentation on current and all previous versions is available at semarchy.com/documentation.

Workflows and Business Processes

xDM supports stewardship using human workflows for data authoring, validation and duplicate records management. Using graphical human workflows, business users, data stewards and data owners can collaborate to contribute to the highest quality content for the golden data records.

Stewardship workflows are designed and documented within the logical model using BPMN, often referred to as "whiteboard modeling." They are automatically deployed when the data model is instantiated. Workflows are version controlled with the model.
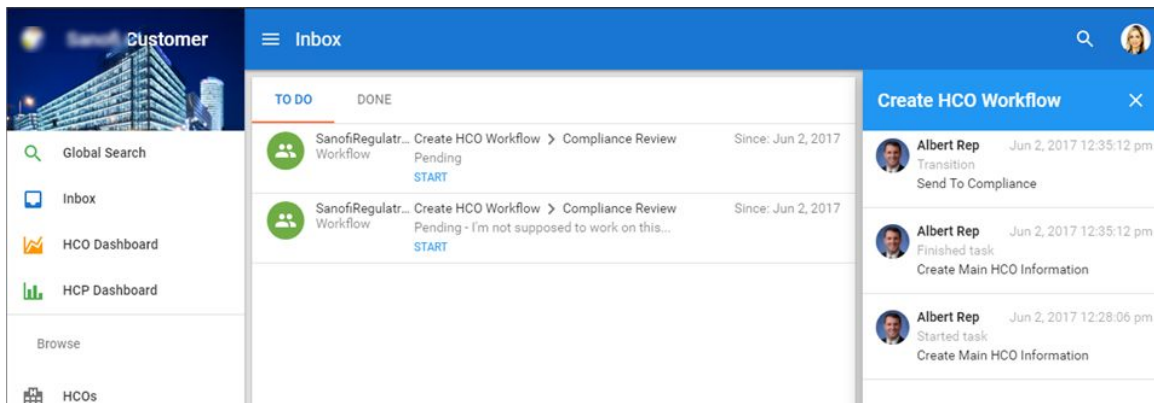
Inbox with tasks assigned for data approval.

The example below shows a workflow between 3 stakeholders: A Sales Rep in charge of creating Healthcare Organization data, a person from the Regulatory Compliance team in charge of reviewing such data with regards to data privacy and compliance rules, and finally a Data Steward in charge of performing the final sanity checks before submitting the approved data to the MDM.
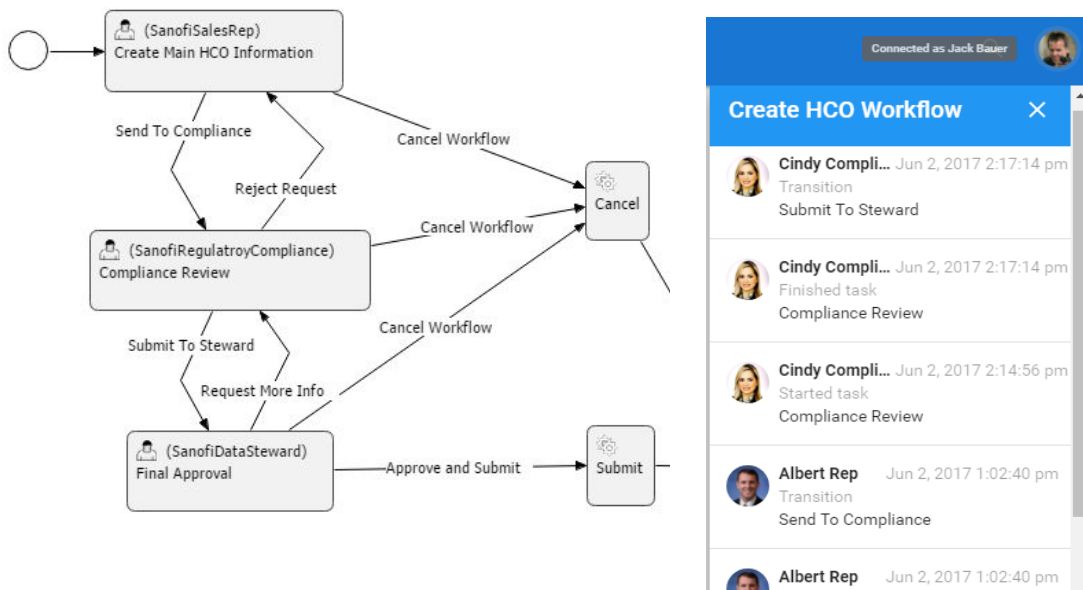
Multi-step approval workflow design (left) and traceability/collaboration at runtime (right)

Once deployed, workflows are triggered using actions (such as Create, Edit, Mass Update, Delete, Import, etc.) available to user roles.

Business users and data stewards access the workflows from their main xDM application inbox or from their emails to take actions for authoring additional data, verifying, approving, or re-assigning their tasks.

Workflows automatically enable, disable, or hide any fields or action depending on the user privileges. Workflow tasks can be customized and data-driven to expose only a subset of the fields according to the context (based on session variables or data values, etc.)

All actions taken by collaborative members of the workflow are logged by xDM, including all data changes.

xDM includes an email notification mechanism to inform workflow participants when their contribution is required. It also exposes badges in user-facing applications to guide users to the actions requiring their attention.
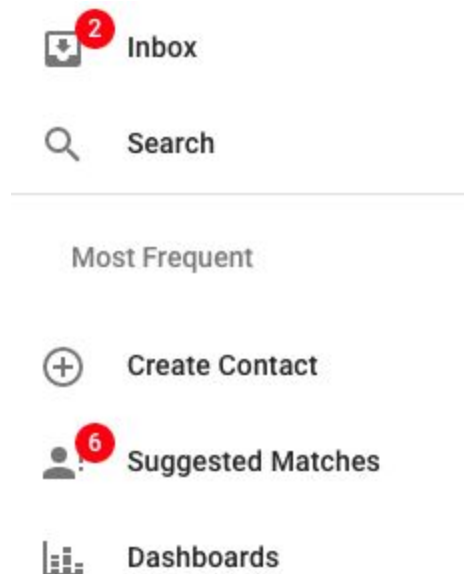
Badges guide business users to required actions.

Performance analytics is performed using **xDM Dashboards**. Standard dashboards allow linking workflows, data quality and business metrics to build unique and fully customizable visualizations, exposing key performance and benchmark indicators for business users and data champions.
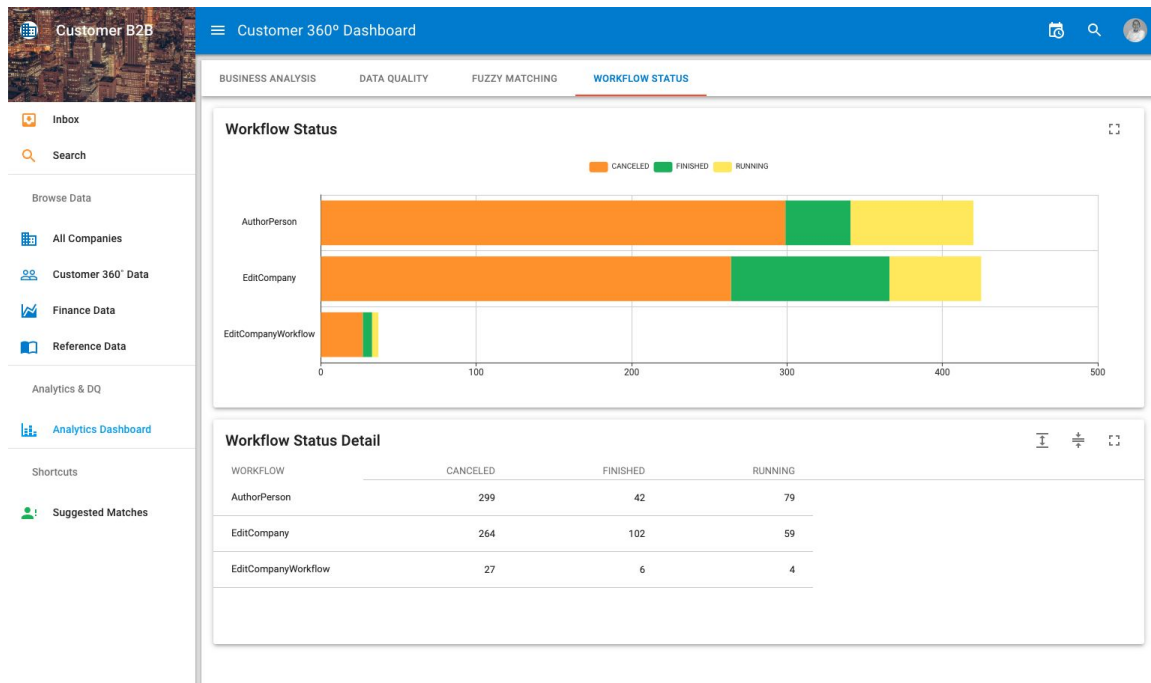
Simple workflow status monitoring dashboard.

Key Capabilities

- Policies/rules defined in the model are automatically enforced in the workflow
- Privileges for data access (at row and attribute level) are automatically enforced to guarantee security compliance
- Tasks can be assigned or re-assigned to roles (groups of users) or individual users
- Rules can be defined to dynamically escalate or re-assign tasks to ensure compliance with SLAs.
- SLAs can be applied to tasks to define a maximum pending duration or a duration to complete. Non-compliance with SLAs can trigger special actions such as escalation, re-assignment, or user-defined hooks.
- Workflows are designed graphically using BPMN.
- Advanced notifications using email, JMS, or other notification mechanisms allow for maximum interaction with users.
- In-application notifications and badges guiding business users to take action.
- Isolated data manipulation transactions can be shared across roles without impacting the golden data of the hub.
- Built-in workflow metrics and KPIs are automatically exposed in xDM Dashboards.

- REST APIs are exposed natively by the platform for potential integration with external business process engines (data, rules, UI Forms, workflows)

Loading/Synchronization/Business Services/Integration

xDM supports both batch integration using ETL/Data Integration products and real-time integration patterns. Both patterns can use either of the built-in SQL or REST APIs.

Designers of the data hub models can define their subscription or publishing endpoints called named queries. These then generate the appropriate APIs (REST or SQL), supporting complex structures as query parameters.

The Semarchy xDM Integration Guide gives further details about the supported integration patterns.
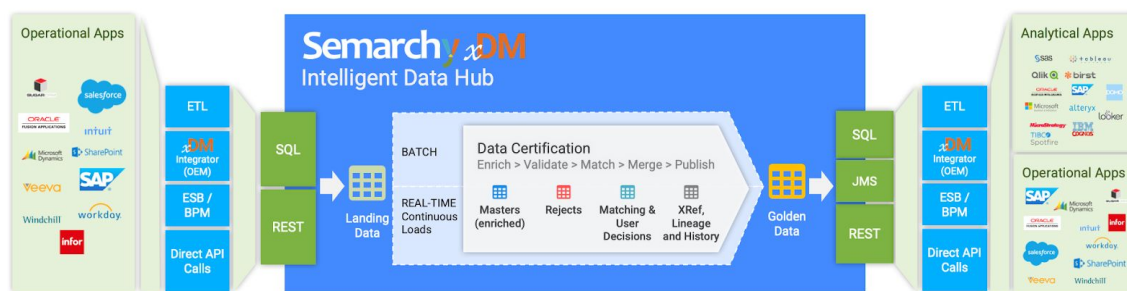
xDM integration pipeline architecture

Batch loading or real-time data submission always flows through the **data certification process** that ensures consistency and accuracy of the received data.
The data certification process is highly optimized for
- **Real-time** access/low-volume per calls/high velocity through the xDM Continuous Loads asynchronous mechanism; and
- **Batch loading**/high volume/low frequency type of integrations using in-database set-based processing.

Both integration patterns share the exact same logical rules for data enrichment, data validation, matching and merging.

Most of our clients use their own integration middleware to integrate with xDM (Informatica Power Center, Oracle Data Integrator, SAP PI, Snap Logic, Jitterbit APIs, Talend, IBM DataStage, MuleSoft, manual coding, etc.). Some of our customers use our OEMed Data Integration solution (xDM Integrator).

ETLs and ESBs benefit from the **SQL API**s as well as the **REST API**s for submitting data to the hub. They can also be used for consuming the data from the hub to populate downstream analytics or any operational applications using any of the **SQL**, **REST,** or **JMS** APIs.

Point-to-point real-time integration for various use cases is usually done through our REST APIs. Those use cases include:

- Search in data hub before simultaneously creating a record in both the operational application and the hub.
- Embedding the data hub records within the operational application UI using configurable queries.
- Portal / process integration for any of the data managed within the hub.
- Data virtualization for building an enterprise virtual layer.

Querying REST APIs for reading or writing data respects the logical model semantics. xDM has built-in auto-documented APIs dynamically generated using swagger standard. This has greatly increased the adoption of the APIs by our customer base.
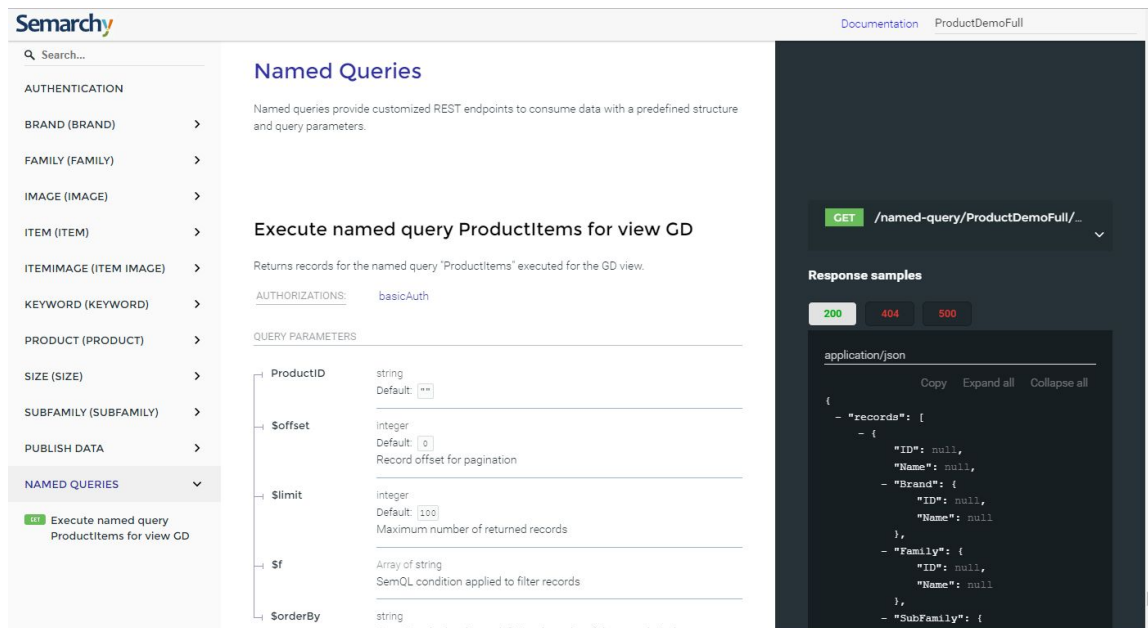
Example of a generated REST API endpoint

Performance benchmarks are regularly executed by Semarchy QA teams in collaboration with our Customer Advisory Board.

Our team has benchmarked a highly loaded environment for a Government related portal with a peak test of **1,800 requests per second** hitting a B2B database of over **20 million golden parties** for 1 hour.

xDM REST APIs have performed as expected with a constant response time of **~48ms.** The architecture was using a PostgreSQL database load balanced on 3 mid-sized application servers (2 CPU Cores, 8GB RAM).

Analysis Summary

Period: 10/09/2018 09:12:30 - 10/09/2018 10:18:27

Project Name:	<%CompanyName%>
Test Name:	<%SessionName%>
Test Description:	<%TestDescription%>
Controller Run Time:	<%RunDate%>
Duration:	1 hour, 5 minutes and 57 seconds.
User Notes:	<%UserNotes%>

Statistics Summary

Maximum Running Vusers:		100
Total Throughput (bytes):	⊘	8 020 676 840
Average Throughput (bytes/second):	⊘	2 026 447
Total Hits:	⊘	6 543 848
Average Hits per Second:	⊘	1 653.322 **View HTTP Responses Summary**

5 Worst Transactions

All transactions were in the SLA boundary

Transaction Summary

Transactions: Total Passed: 6 531 041 Total Failed: 12 807 Total Stopped: 0   **Average Response Time**

	Pass			Fail		Stop		
Total	6 531 041			12 807		0		

| Transaction Name	SLA Status	Minimum	Average	Maximum	Std. Deviation	90 Percent	Pass	Fail	Stop
TR2_WS_CONTROLERSIRET	✓	0.006	0.048	0.88	0.007	0.054	6 531 041	12 807	0

Service Level Agreement Legend: ✓ Pass   ☒ Fail   ⊘ No Data

Extract from the benchmarking test in a highly loaded real-time environment

Key Capabilities

- Built-in publish / subscribe framework with real-time data notifications.
- Support of SQL, REST and JMS APIs with shared authentication and security policies for data access and management.

- REST OpenAPI description and documentation with multi-language client generation using swagger.
- Massive bulk loads are highly optimized as Set-based SQL processing inside the RDBMS engine.
- Same optimized data certification processes executed regardless of the integration pattern (ETL/Batch, ESB/Message-driven or REAL-TIME/Point-to-point) to guarantee data consistency and accuracy.
- Integration with existing ETLs/ESBs and BPM software is greatly facilitated.
- Possibility of using xDM Integrator – Optional OEMed component (see screenshot below) – as the data movement / transformation layer to capture data from any sources and populate any targets.
- Native support for Registry, Consolidation, Coexistence, and Centralized styles within the data hub.
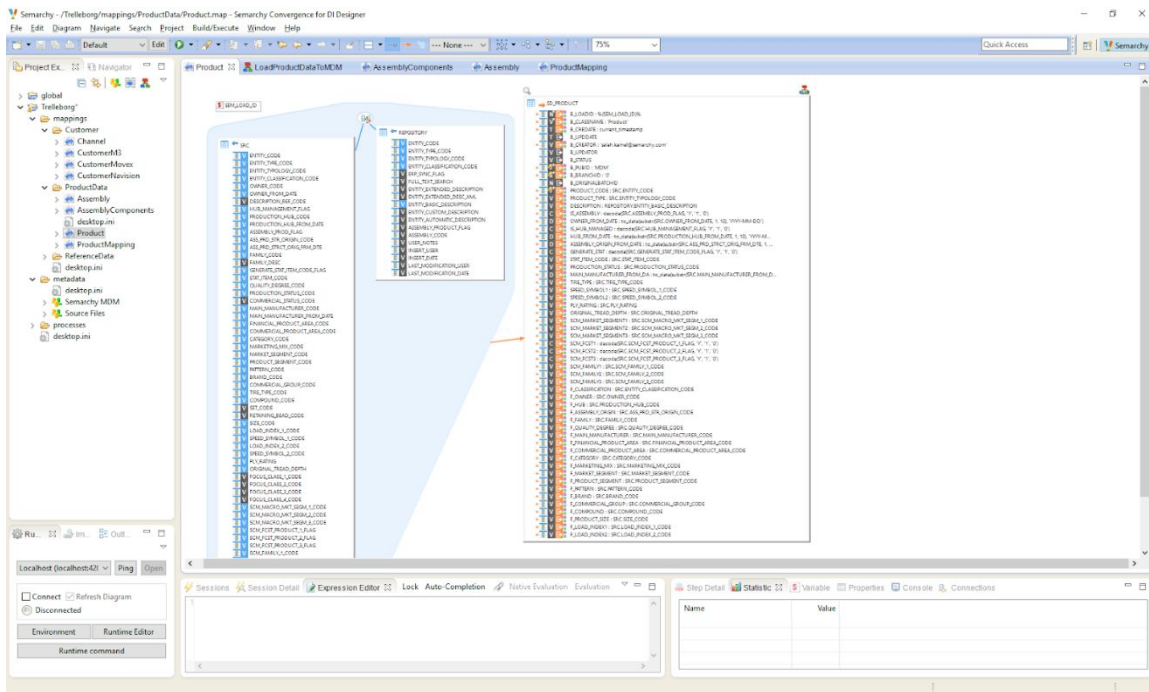
A source to data hub integration mapping in xDM Integrator

Data Modeling

xDM provides a fast and iterative modeling framework that simplifies the design of semantically complete enterprise data models. It enables collaboration between business users, data stewards, data champions and data architects for defining high-level concepts expressed using business terminology.

xDM Data Models act as the single point of logical definition of entities, attributes, relationships, and rules managed in the data hub.
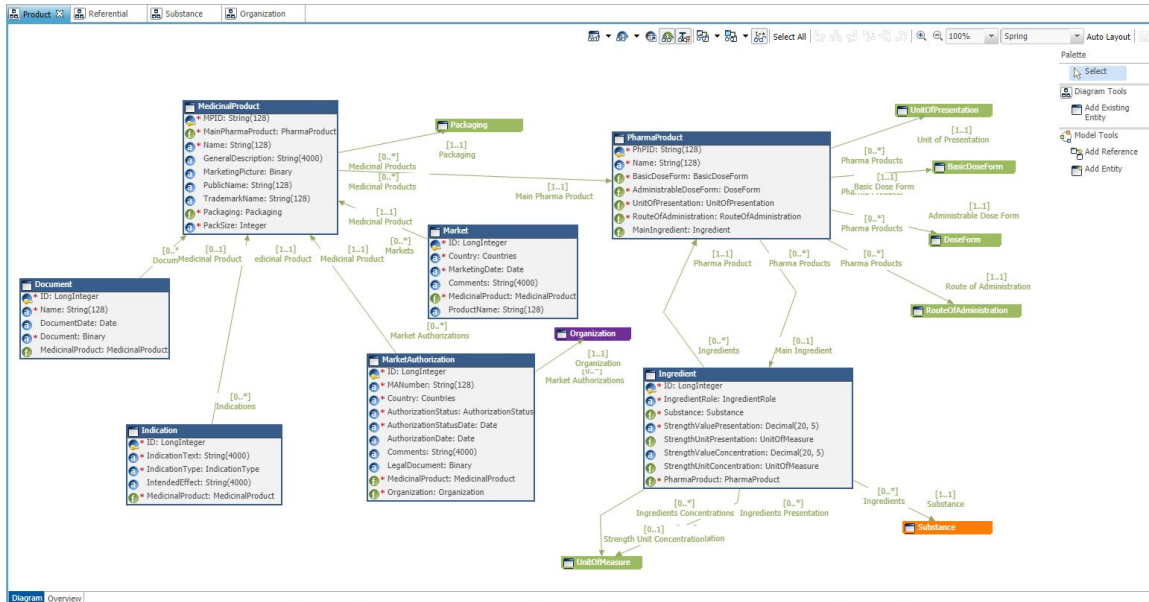
Example of a subset of an IDMP data model in xDM Application Builder

Once a version of a data model is completed, it is deployed by the platform to the underlying database. Several versions of the same model can be deployed at the same time and xDM ensures forward and backward compatibility between the various data model versions as deployed by the data architects.

The data modeling flexibility at design-time and at deployment time is one of the most important differentiators of the platform. It is the key driver that enables the agility to support rapid implementation of changing business requirements.

Support of Centralized, Consolidation, Coexistence and Registry styles is expressed within the data model entities and can be changed over time.

Centralized / Registry styles usually use the "*Basic*" entity type.

Consolidation and Coexistence styles will prefer the "*Fuzzy Matched* "or "*ID Matched*" entity types.

All entities of the data model support bi-temporal modeling (Historize Golden / Master Data) and automatically generate the appropriate **historization** and **lineage** data infrastructure and processes.
Entities also support record deletion in all hub styles (including consolidation), handling cascade and veto through the model's relationships; Similarly to historization, the processes, structures and APIs required to manage and propagate deletions are automatically generated.

The data model, rules and policies are exposed to users within the data management application as documentation endpoints, supporting full navigation as well as contextual help, putting data literacy at the core of the data hub.
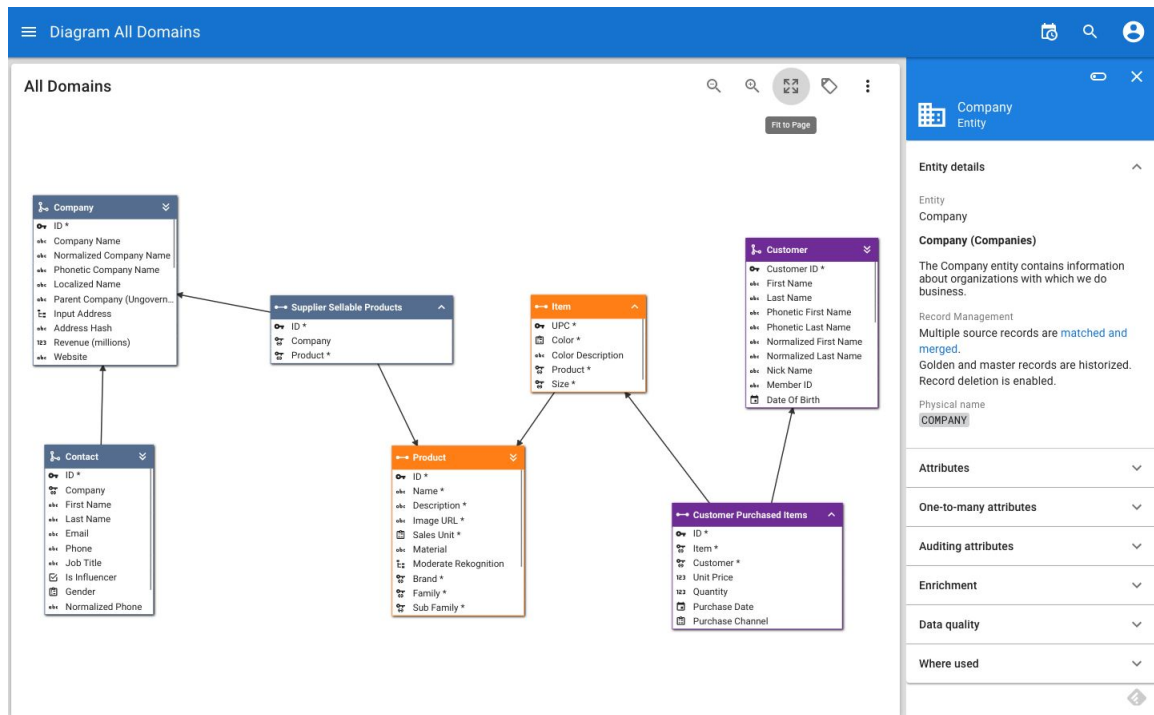
A model diagram exposed in the application documentation

Key Capabilities

- Full **multi-domain modeling** framework supporting all end-user requirements.
- Support of reference data, master data and transactional/observational data modeling patterns.
- Concurrent support for **all hub styles** (registry, consolidation, coexistence and centralized)
- Support of **graph-based** modeling for complex relationship resolution. Refer to the "Why MDM & Graph?" presentation for more details.

- Data Model includes entities, attributes, relationships, complex types, inheritance data rules, policies, data sources, user interfaces and workflows in a **single unique graphical design-time** experience.
- **Data rules** rely on a powerful graph language (**SemQL**) and/or an extensible plugin framework for defining
  - In-line **data enrichment** and data standardization rules.
  - **ML**-based rules for dynamic **data classification**.
  - External **AI-driven microservices** call-outs (Google, Amazon, Microsoft, etc.)
  - **Data validation** and **data quality** rules.
  - **Match** rules with scoring.
  - **Merge/consolidation** rules with attribute value lifetime definition
  - **Security** and **privileges** policies.
- Simple user experience for defining and enforcing **horizontal** (record-level) and **vertical** (attribute-level) **security** policies.
- Built-in **internationalization** and localization of all the elements of the data model for WW deployments.
- Automatically generated **business-consumable documentation** and link to the data governance **glossary**, **processes**, **policies**, **domains**, and **catalog**.

Information Quality and Semantics

xDM provides an end-to-end solution for data quality and semantics. It includes the following modules out-of-the-box:
- **xDM Discovery**: main new component to automatically profile any data source
- **xDM Data Certification Process**: core component for data quality rules implementation and enforcement (enrichment, complex validations, match/merge)
- **xDM Dashboards**: core component for continuous data quality monitoring, analyzing master or transactional data, and for data quality KPI definitions

xDM Discovery - Data Profiling

xDM Discovery enables data architects and business users to gather metrics and profile any source data to prepare a data management initiative or to maintain data governance KPIs over time.

xDM Discovery connects to data sources, profiles the existing data and exposes these profiles in the xDM repository. Dashboards are auto-generated for core metrics. Profiles include high-level table metrics and detailed column metrics, such as
- lowest/highest and most frequent value

- uniqueness
- nulls
- minimum, maximum and average value
- length value
- pattern distribution
- dynamic remediation / usage recommendations on each profiled column
- related columns

Business users and data champions analyze the profiles using the fully automated, built-in xDM Discovery dashboards with drill down to the source data to identify root causes and/or potential remediation processes.

Advanced profiling is also possible with a fully configurable dashboarding experience using **xDM Dashboards.**
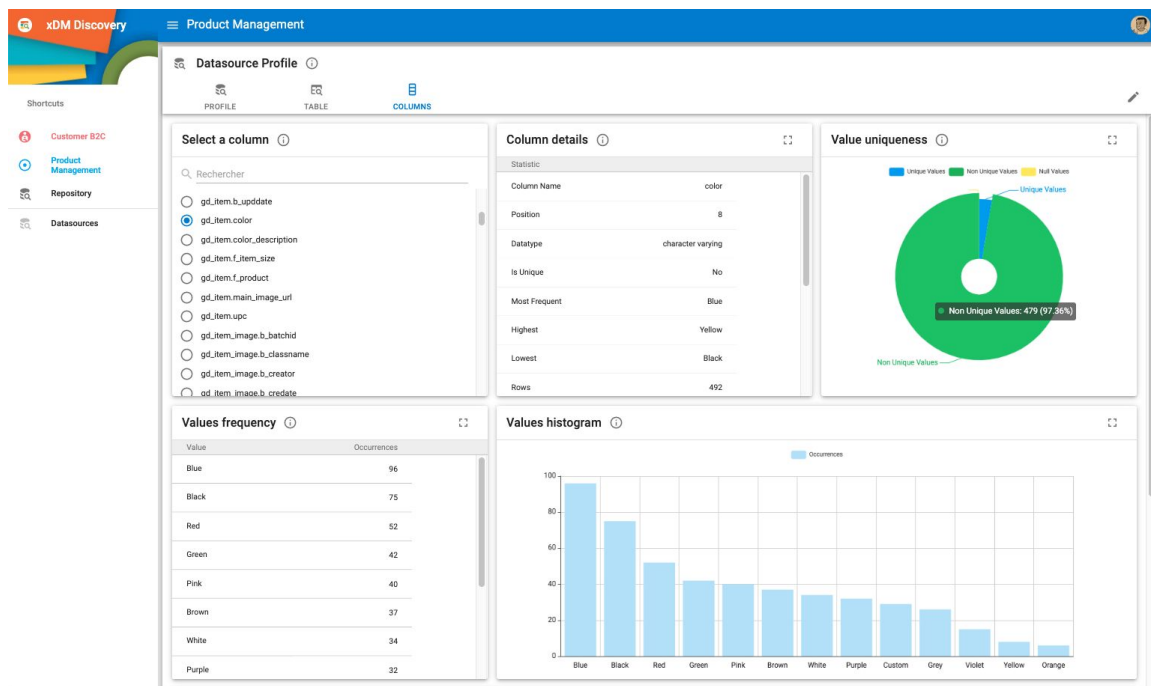
A built-in profiling dashboard with column statistics and values distribution.

With data discovery metrics, data champions have a clear assessment of the data quality at any time and measure the evolution of such metrics against KPIs over time.

In an MDM centric environment, xDM Discovery can, for example, be used to:
- Understand the ad-hoc data quality of a source before onboarding it into the Data Hub;

- Monitor the completeness, consistency and accuracy of a set of attributes in the Data Hub;
- Identify related attributes through their data values and derive an impact analysis;
- Find redundant or derived data elements in order plan for optimization of operational data authoring costs;
- Derive data quality rules and enforce them using xDM;

**xDM Data Certification Process - Data Quality Rules Enforcement**

The **xDM Data Certification Process** transforms, enriches, standardizes, validates, matches, links, and reconciles records provided to the data hub by source applications or authored by business users using the data management UIs or APIs. The result of this data certification is then exposed as certified golden data.

The automated Data Certification Process is entirely generated from the rules defined in the data model for both batch and real-time certification. The process oversees creating and maintaining golden records over time, whether these records were changed in the source, updated by a user in the xDM application, automatically/manually merged, or manually unmerged. Data lineage and audit trails are exposed for compliance requirements.

Enrichment / Standardization Rules

xDM Data Certification Process uses the data enrichment rules to augment, cleanse, or standardize the data using trusted external data providers or automated data transformations.
Enrichment rules can use Semarchy data manipulation language (SemQL) that allows for graph relationship resolution, data transformation and complex nodes filtering with a built-in library of over 150 functions.
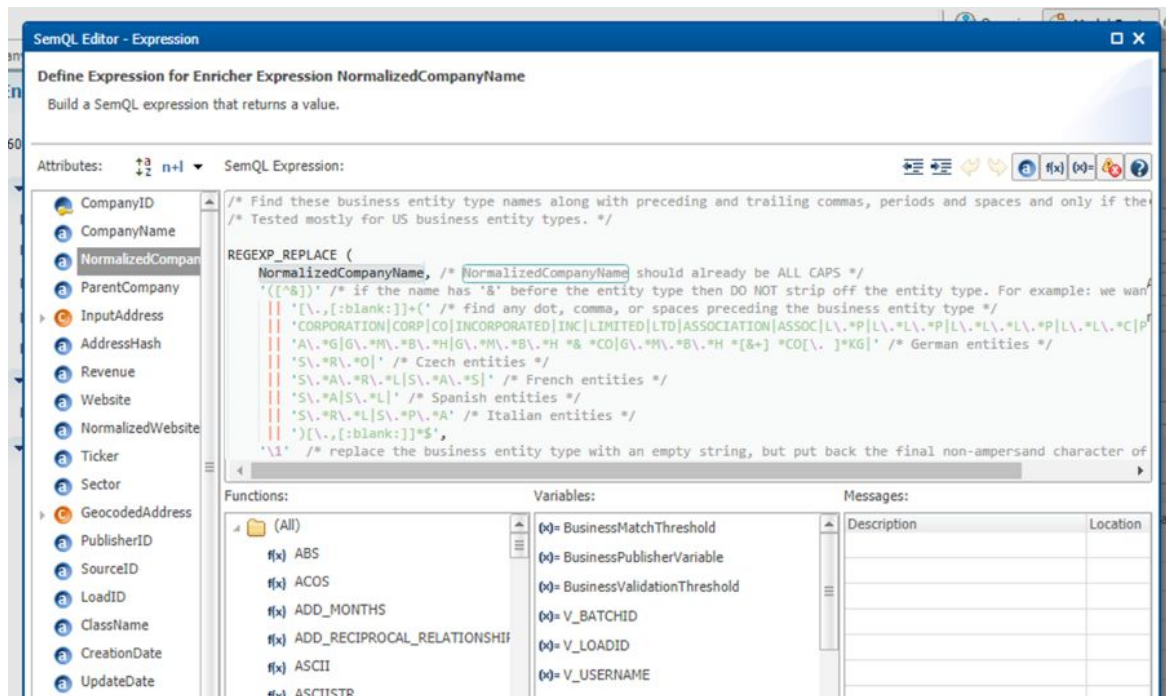
Example of a data standardization rule using SemQL regular expressions

Typical use cases for SemQL include basic data cleansing such as substring, regular expression text parsing, case conversion, if/then/else logic, complex calculations, value lookup, Soundex and other Metaphone algorithms, etc.

xDM also provides a plugin framework to integrate with **external data-/micro-services** providers.

Example of the Amazon Image Recognition plug-in configuration

Enrichment plugins fit in the following categories:
- **Content-lookup enrichment plugins**: B2B Party lookup (Experian, D&B), Address lookup (Google, Bing, OpenStreetMap), Person name (with gender/nickname detection), Phone Standardization (locations, carrier, time zones phone types, etc.), Email with DNS lookup/auto-correct, text parsing and advanced text phonetic transformation (Metaphone, double-Metaphone, Caverphone, etc.), language transliteration, transcodification using governed data dictionaries, etc.

- **AI/ML-based enrichments**: Amazon Rekognition for images parsing, Sentiment Analysis, Text extraction with domain-specific ontologies, OCR leveraging USCDataScience, Dynamic classification of data based on textual input and taxonomies, Language translation using Google Translate, etc.

- **REST Clients** invoke a REST service API that is exposed for communication by any system or service provider such as Google Geocoding API, or Melissa Global

Email Web API. Rest clients are designed in Semarchy xDM with no code and used as an enrichment as for B2B Party lookup (Experian, D&B), Address lookup (Google, Bing, OpenStreetMap), Person name (with gender/nickname detection), Phone Standardization (locations, carrier, time zones phone types, etc.), Email with DNS lookup/auto-correct
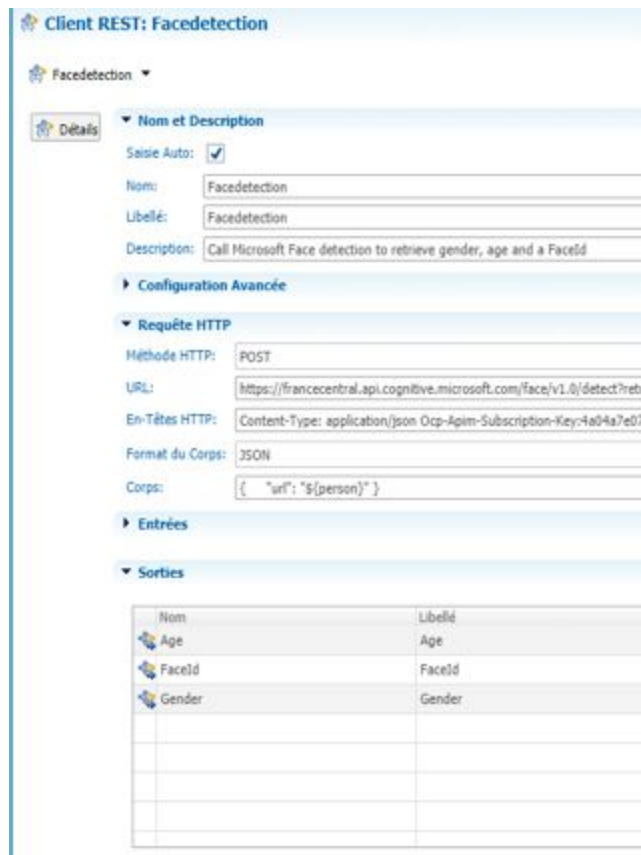
Example of the Microsoft Face Detection Rest Client configuration
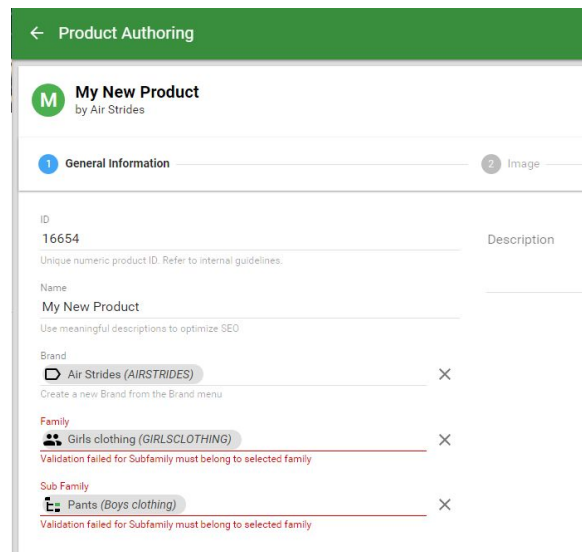
Validation Rules

After the data is enriched by the xDM Data Certification Process, it is validated against the Data Validation Rules. These rules follow the same logic as the Data Enrichment Rules where data architects can choose between SemQL and Plug-ins to define the acceptance criteria for records in the data hub.
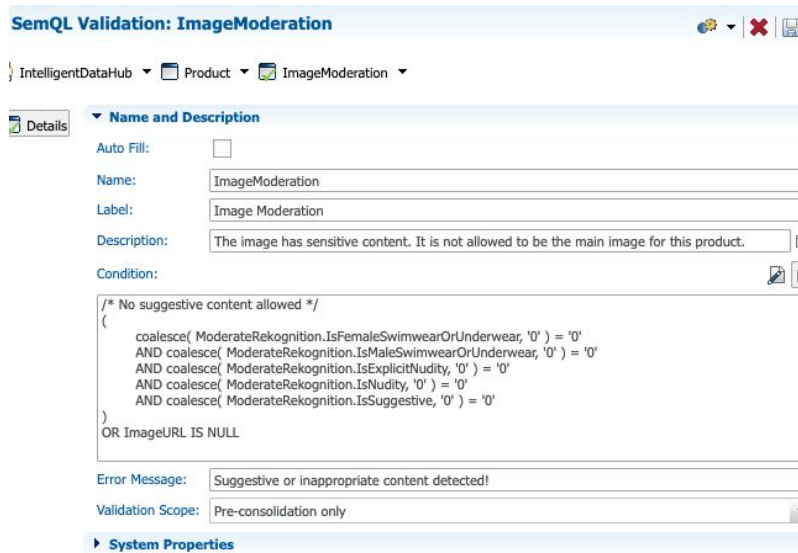
Example of a data validation rule in xDM

The xDM Data Certification Process automatically rejects the records that do not meet the data quality requirement as exposed by the rules. Rejected records are then pushed into work queues proposed to business users or data champions for manual review and remediation.

During data authoring, the same data enrichments and validation rules are triggered in real-time in the guided data authoring UI forms to provide immediate user-customized feedback while manipulating the data.

Example of a real-time data validation in the data authoring UI

Match Rules

xDM provides a built-in [Fuzzy Matching](#) engine leveraging the power of SemQL to detect arbitrary matches between pairs of records. Those pairs of records are automatically aggregated into match groups that will later form unique golden records.


Example of match rules definition in xDM

Multiple Match Rules in a Matcher (see example above) define several conditions for considering two records a match with scores that represent the percentage of confidence of the match.

Depending on such a score and the merge policies (thresholds), the matcher automatically merges groups of records together to create golden records.

The match rules outputs are tracked and exposed to data stewards as an "explain data" graph view.
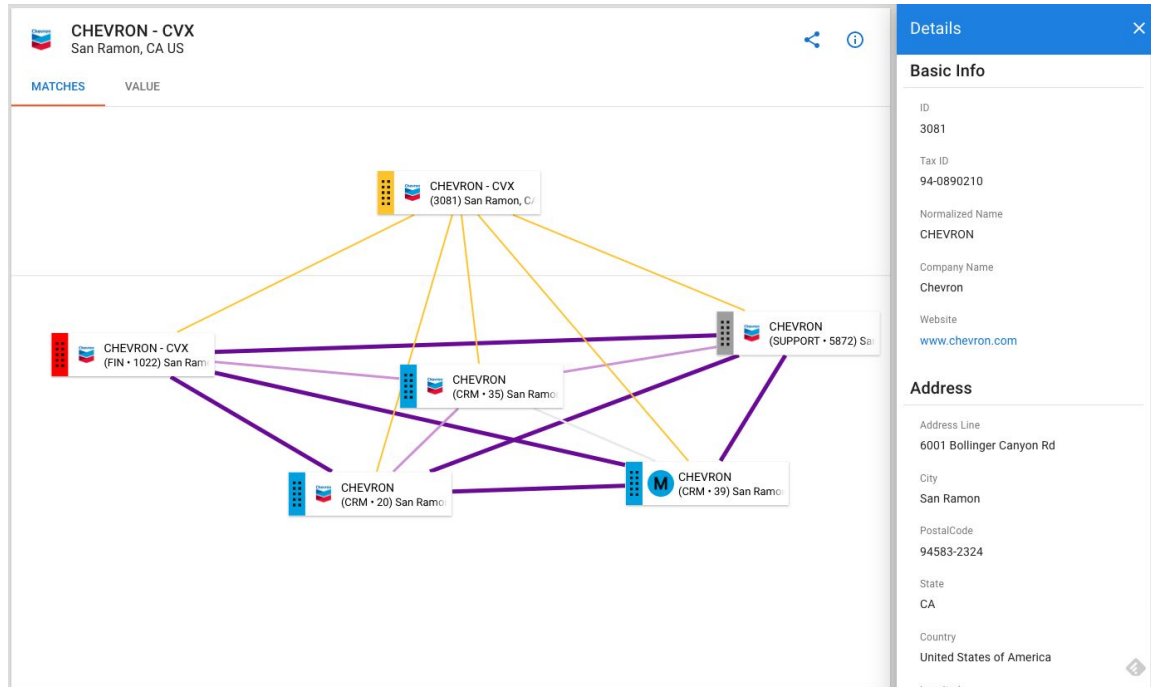
Data champion UX to analyze matches and golden clusters

When a group is not merged automatically, because of a low score, it is flagged as a Suggestion and queued for further screening by data champions. They can use advanced duplicate management actions to decide whether to merge or unmerge groups and create or split golden records.
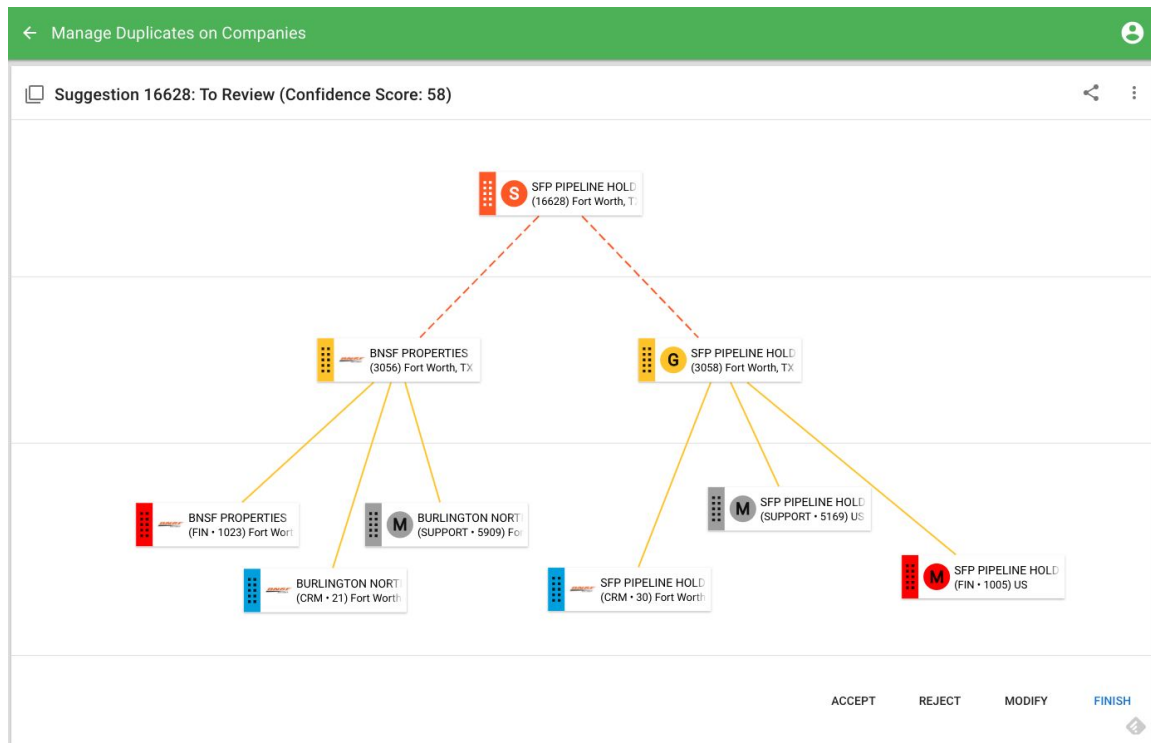
Data champion UX to review duplicate management suggestions

Semarchy xDM version 5.2 introduced major additions to the match and merge engine to align more closely with the intuitive expectations of data stewards and business users.

- **Transitive Scoring** is a new algorithm to compute the similarity score of a group of matched records. It treats the score of each pair of matched records as a probability and combines these probabilities for all possible paths in the group. Using this algorithm, designers can define more accurate scores and thresholds for the groups without modifying existing match rules, resulting in more reliable similarity clusters, and requiring less human curation.
- **Multi-Iteration Grouping** is another evolution of the match and merge engine. It uses a multi-pass algorithm to iteratively create subgroups with a high similarity score within larger groups with a lower similarity score. Designers can choose to turn this option with a simple checkbox resulting in less, but more accurate suggestions for the data stewards, and therefore reducing the time and effort spent in manual tasks.

Survivorship (Merge) Rules

The quality of golden records depends on the ability of the data hub to consolidate various pieces of information from multiple sources whilst picking the most accurate values (from a business standpoint).

[Survivorship Rules](#) are defined collaboratively in xDM by data architects, data champions and business users. They indicate which data survives in the golden records per attribute as well as the lifecycle of the user overrides when those values are edited/modified in the hub.

Consolidation rules expose multiple strategies for ranking values across sources, including **recency**, **frequency**, **completeness**, **preferred sources**, and any **custom ranking** algorithms using SemQL.

Managing the lifecycle of the data of the overrides is made easy with multiple pre-built strategies to define when the user overrides expire.

Golden record with survivorship rules and value highlighting

Deletion

Semarchy xDM allows deletion of golden and master data with propagation/veto rules to comply with GDPR/CCPA and other privacy regulations.

A delete action may be:
- A Hard Delete that physically deletes records from the hub database. With such a delete, data cannot be recovered.
- A Soft Delete that logically deletes records by moving them to a deleted records storage.

Traceability and Historization

Semarchy xDM has an automated mechanism to trace all data changes and historize golden and master data for all entities. Model designers can configure historization for new or existing entities, at the click of a button, in order to trace every record change, creation, update or deletion. Semarchy xDM automatically creates the structures and processes to store and maintain the record history.

Once historization is configured, MDM Applications allow business users to browse the hub at any point in the past, viewing data as it was before. They can also review the full history of changes for a given record. Record history is also available for integration via the built-in REST API.
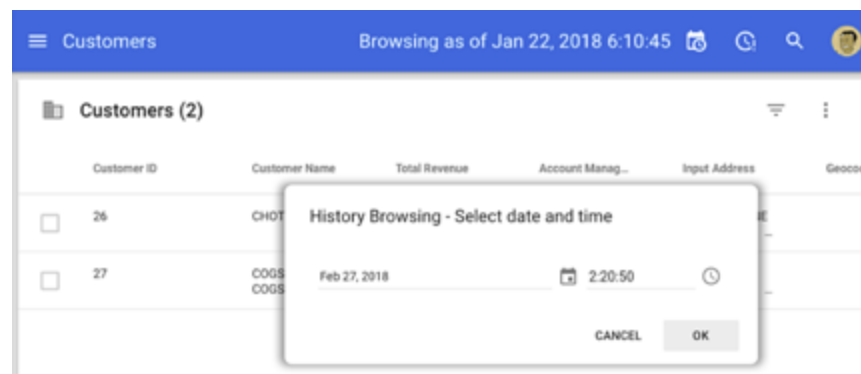
Business users can easily start browsing the hub at a given point in time

Using this feature, data architects can use the MDM hub as a safe storage for all data changes, and as a single point of truth for data traceability, available for both applications and business users. The MDM hub becomes a central point in the enterprise data architecture where the history of all critical data assets is preserved for compliance purposes.

xDM Dashboards - Data Quality Monitoring & Contextualization

Data quality metadata (definition) and metrics (enforcement) are stored by default in the data hub. The performance of the data quality of the hub can be monitored in the platform at any point in time.

xDM Dashboards is the component of the platform for designing visualizations and dashboards that create insights into the data quality metrics, possibly in context with business data and the data stored and managed in the data hub.

Dashboards may be exposed within the xDM Data Management applications or in dedicated Dashboard Applications.
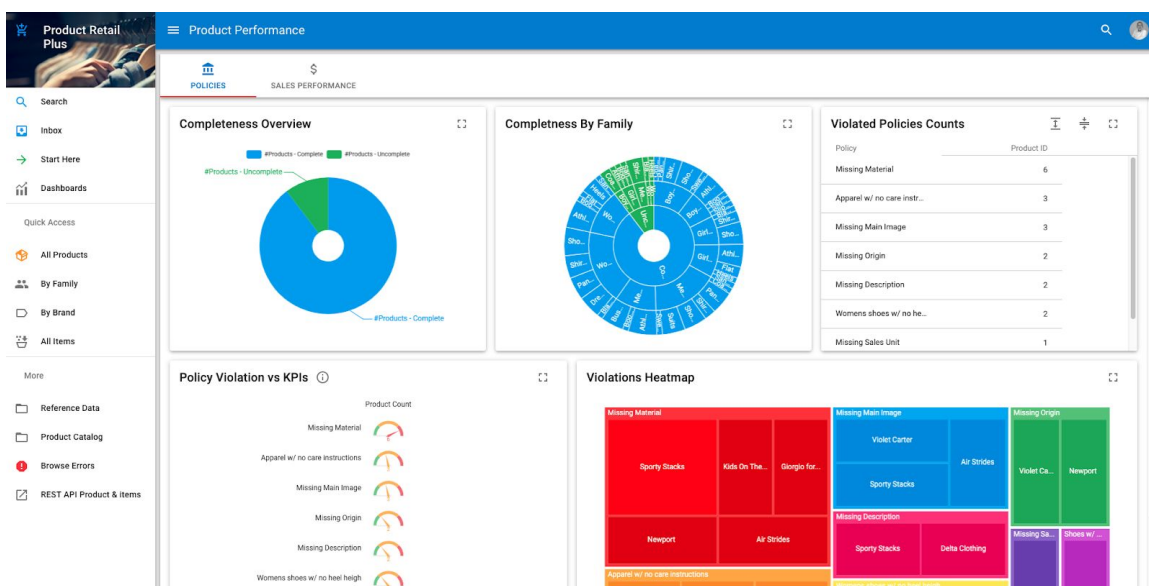
A product data quality dashboard built with xDM Dashboards and embedded in an xDM data management application

The data quality metrics stored in the data hub are also exposed via APIs and may be consumed by any external dashboarding/data visualization products. Such third party dashboards can be integrated seamlessly with the xDM Data Management applications.
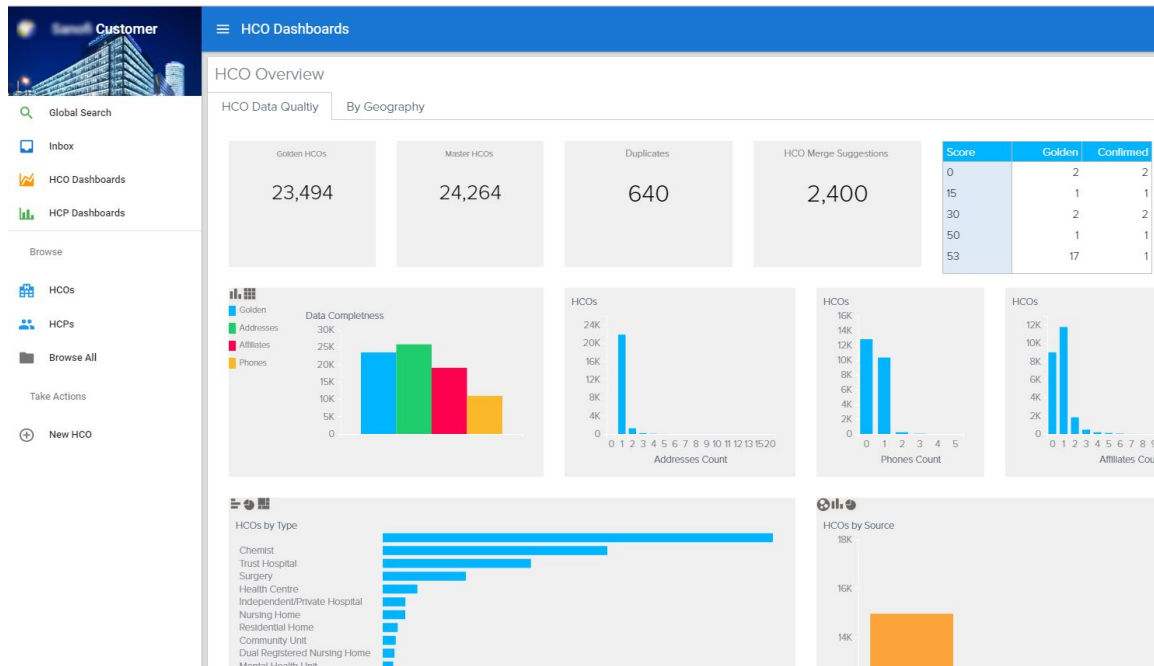
A data quality dashboard, made with a third party data visualization product and embedded in an xDM data management application

Performance/scalability, Availability and Security

Performance/scalability and Availability

xDM provides an architecture for enterprise-scale deployments, based on Java EE. xDM uses **PostgreSQL**, **Microsoft SQL Server** or **Oracle** relational databases for storing and enforcing the CRUD consistency of the hub data. The relational database engines optimize the data certification process, standardizations/enrichments, data quality validations, and the match and merge algorithms.

The certification also supports built-in features such as caching, multi-threading, and batch processing to speed up the certification process, reducing the possible lags caused by the use of external data providers and services.

The platform is designed, tested, and certified for clustering, load balancing, enterprise security, fault-tolerance, and high-availability configurations, both on-premises and in cloud deployments. *Refer to section Loading/Synchronization/Business Services/Integration section for real-time scalability benchmarks of the data access REST API.*

Semarchy clients include several very large-scale deployments. Typical configurations of our clients are summarized below:

	Small	Medium	Large
Total number of records in the hub	< 10M	10-100M	100M+
Number of users	Up to 500	Up to 5,000	5,000+
Recommended RDBMS (on-prem)	Oracle Standard Edition, PostgreSQL or SQL Server Standard Edition 1-2 CPU/8GB RAM	Oracle Enterprise Edition, PostgreSQL or SQL Server Enterprise Edition 2-4 CPUs/16-64GB RAM	Oracle RAC, PostgreSQL or SQL Server Enterprise Edition
Recommended RDBMS (cloud)	Amazon RDS db.m3.medium (Aurora, Oracle, SQL Server or PostgreSQL)	Amazon RDS db.m4.large (Aurora, Oracle, SQL Server or PostgreSQL)	Amazon RDS db.m4.4xlarge + (Aurora, Oracle, SQL Server or PostgreSQL)
Typical Application server configuration (per node) whether on-prem or cloud	Apache Tomcat 8GB RAM, 2CPU	Apache Tomcat 8GB RAM, 2CPU	Apache Tomcat 8-16GB RAM, 2CPU
Recommended Number of app server nodes (in the cluster)	2	2-6	6+

xDM deployments are horizontally scalable by increasing the number of application server nodes in the cluster.
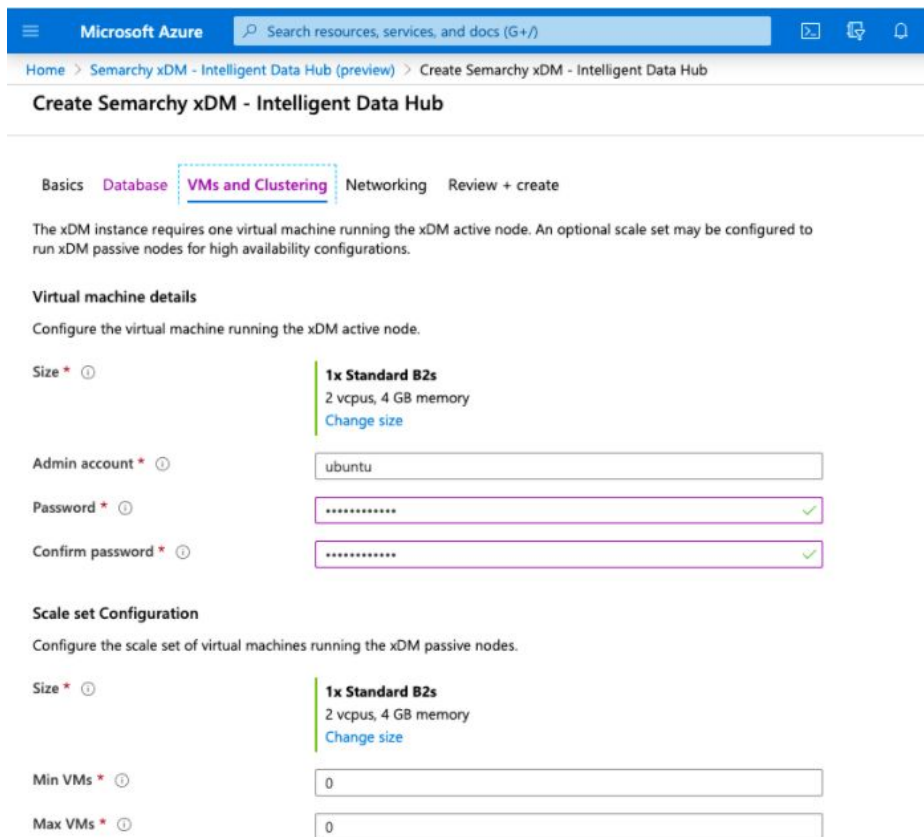
Cluster and sizing configuration for an xDM instance in Microsoft Azure

The database layer is scalable by increasing the number of CPUs and RAM (on-premises) or by extending the type of hardware cluster when running in the cloud. High availability of the database layer uses vendor-supplied standard mechanisms (Oracle HA, PostgreSQL HA, Amazon RDS Multi-AZ, etc.)

Security

xDM uses role-based as well as persona-based security and privilege grants for accessing the application features as well as the data. Fine-grained privileges (attribute/record level – horizontal/vertical security partitioning) are defined as part of the model design and enforced for any data access method.

xDM user interfaces, workflows and REST APIs generated from the logical model automatically apply security policies, hiding fields with no access rights, locking fields with no write privileges, preventing actions such as data exports, etc. In addition,

applications are built as part of the model to export different views, workflows and capabilities depending on the connected users' roles and the usage context.

Hierarchy Management

The xDM logical model defines all relationships between entities and supports an unlimited number of customizable balanced, unbalanced and recursive hierarchies for all data domains. Hierarchies defined in xDM organize entities into single or multiple related schemes for data management, classification, and reporting.
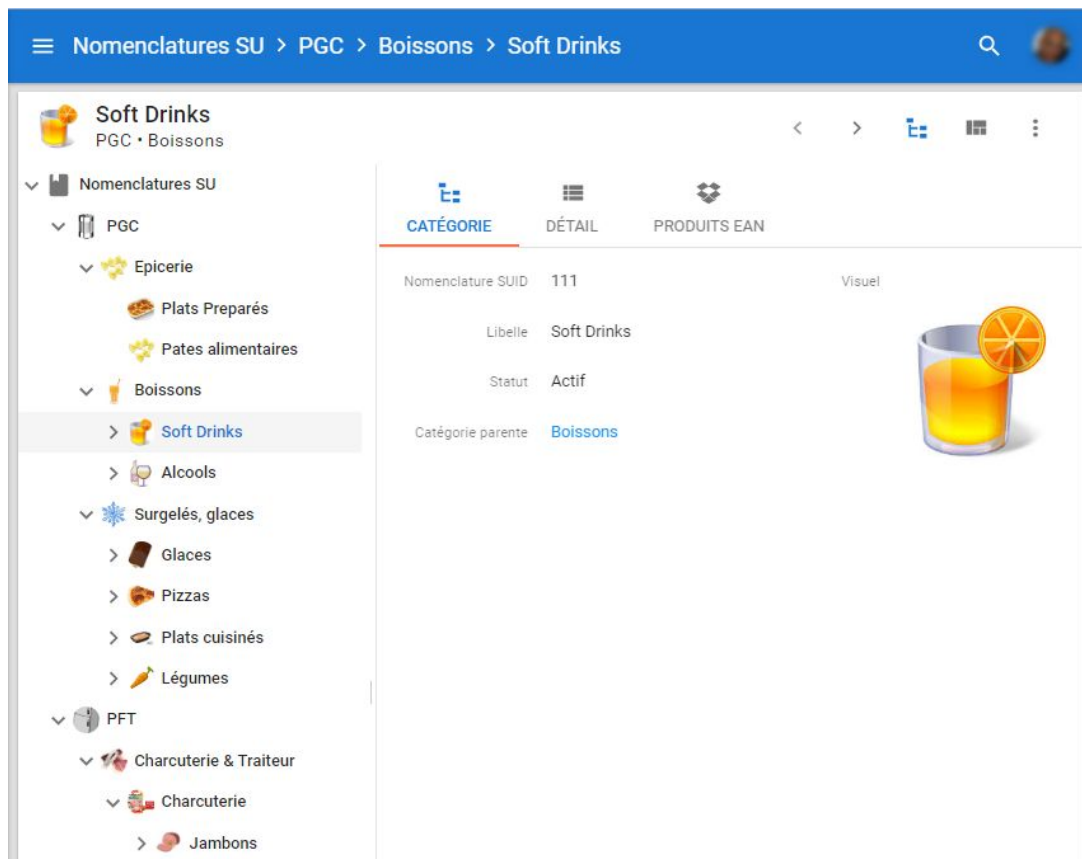
Example of a product taxonomy hierarchy

Hierarchies defined in the model are used for:
- Searching/navigating, both in the generated user interfaces and via APIs
- Enforcing classification, data quality rules and functional dependencies
- Collaboration for defining new hierarchies and/or mapping to existing hierarchies
- Managing time-varying relationships and browsing hierarchies as-of-date.

The SemQL language provides intuitive hierarchy navigation and search, allowing users to navigate both hierarchical and non-hierarchical relations using the same language patterns.

Information Stewardship and Policy Enforcement

Information Stewardship

xDM provides a flexible and complete framework for generating a Material Design user experience for collaborating and managing the data of the hub. Applications generated by xDM are contextually adapted and optimized for various user personas:

Data Consumer
- Interested in getting the most up-to-date data for a domain.
- Expects data to be clean, consistent, and accurate.
- Driven by a business requirement to quickly obtain data-driven answers.
- Interested in receiving daily or weekly updates about data changes.

Data Contributor
- Wants to quickly enter data in the hub.
- Understands the interest of having centralized data.
- Suffers from the Excel Spreadsheet hazardous effects.
- Willing to collaborate with his peers.
- Wants to keep ownership of the data

Data Steward
- Wants to enforce data quality rules defined by the business to ensure data consistency, completeness, and compliance
- Acts as the gatekeeper of clean data.
- Expects metrics to measure performance of the processes and information quality.
- Spots and resolves data discrepancies with the help of the business stakeholders.

Data Champion
- Understands the impact of data changes across multiple business lines and applications.
- Performs data analysis using graph queries to reveal new relationships and opportunities.
- Understands and measures costs and risks related to bad data.
- Drives the data management roadmap whilst keeping a strong handle on the data.
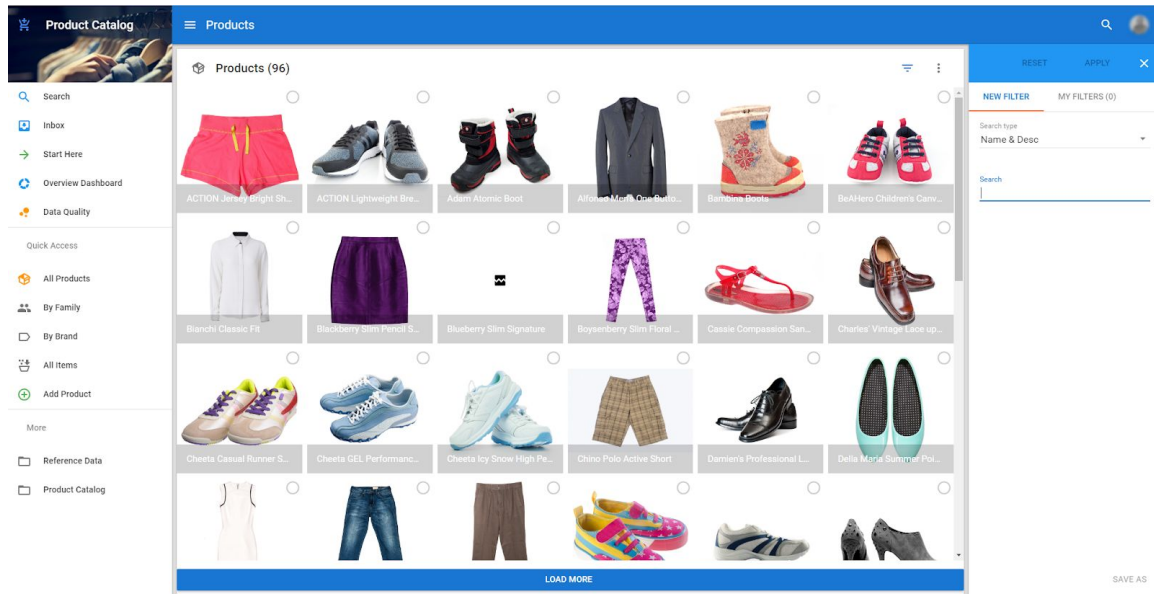
Example of a product data management application

Information stewardship in xDM supports a variety of use cases including:
- **Guided data authoring**, import & export using a step-by-step user experience.
- **Data quality** feedback according to policies and rules.
- **Mass updates** in separate transactions to execute what-if scenarios.
- **Collaborative workflows** for complex data authoring with parallel teams collaborating on the same objects.
- **Work queues** for fixing rejected data.
- **Dynamic data classification** rules with user feedback loop.
- **Review** of automated match/merge and survivorship results using graph view and side-by-side record comparison.
- Manual match/merge with decisions with possible data overrides.
- **Filtering, searching, taking actions** on the data.
- **Graph analysis** for discovering new relationships and opportunities.
- **Inbox** for receiving workflow notifications and data subscriptions.
- Comprehensive linking to the enterprise **business glossary**.
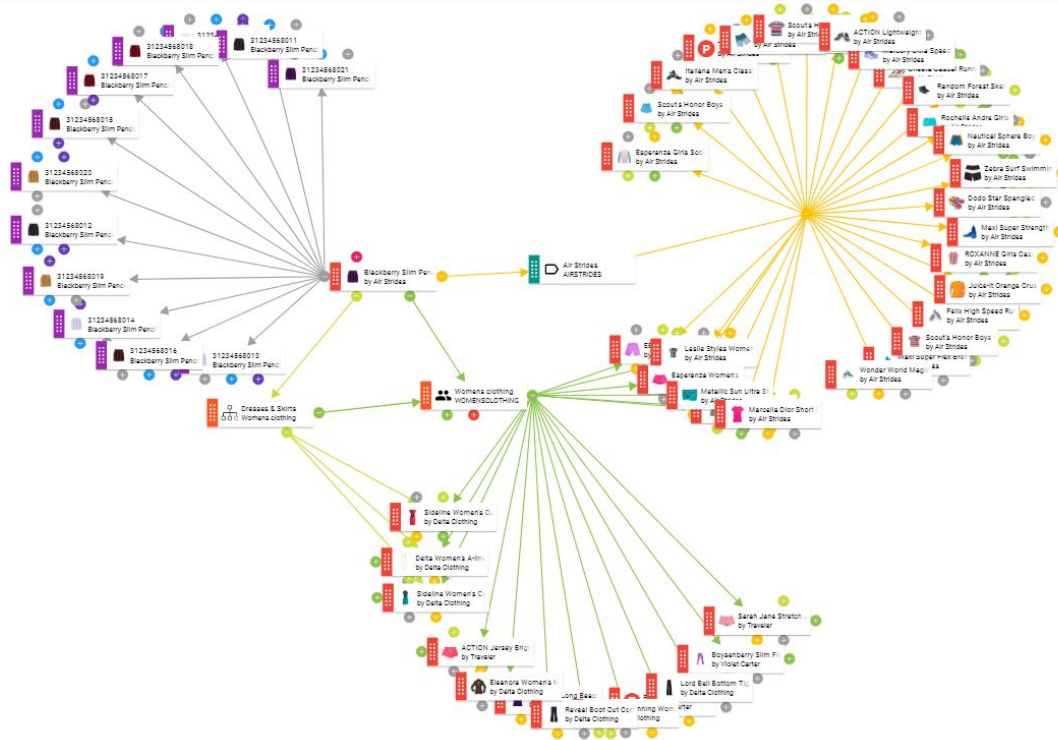
Example of a graph neighborhood analysis in xDM

Policy Enforcement

Security policies, data validation rules, data standardization rules, matching rules, and survivorship rules defined in the logical model (and/or in the Data Governance Application) are enforced automatically for any data managed by xDM (refer to Information Quality and Semantics). Data published via direct authoring or from source applications is automatically enriched, standardized, validated, matched and merged into golden records. Programmatic records consumption or manipulation in the hub is secured according to the data privileges grants defined in the model at entity, attribute or record-level and is subject to the same validation, standardization, and match-merge rules.

xDM guarantees the enforcement of policies (data rules and security/access rules) in the generated data applications, dashboards, and in the REST and SQL APIs.

xDM Dashboard - Built-in Data Hub Analytics and Beyond

The hub performance is tracked using **xDM Dashboards**. In xDM Dashboards, data management operations are monitored in real-time, and every aspect of the data processing and lineage is made available in visualizations contextually adapted for the various personas (data stewards, business users, etc.).

The dashboards are exposed in dedicated Dashboard Applications, embedded and contextualized in xDM Data Management applications or in the corporate portal.

These dashboards can combine a variety of metrics from the data hub, including data volume, lineage, processing efficiency, duplicates management, data quality violations, workflow performance. They can also combine business metrics from external data sources (the operational datastores or the data warehouse, for example).

Dashboards support drill-down to access the un-aggregated data, and support actions to trigger operations in xDM (or external) applications directly from the dashboards.
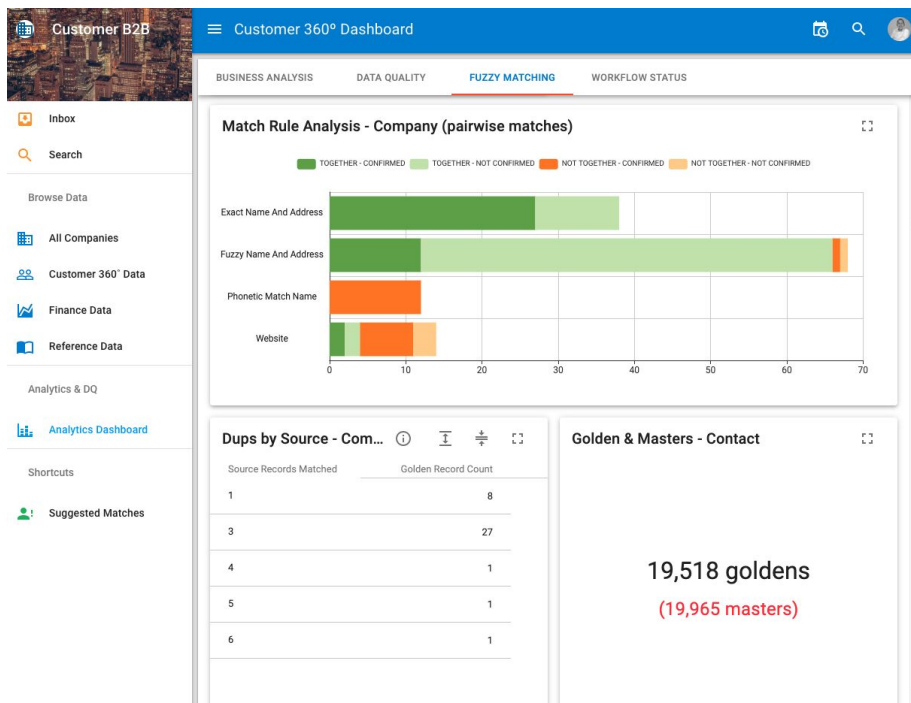
A data stewardship dashboard for fuzzy matched duplicates monitoring
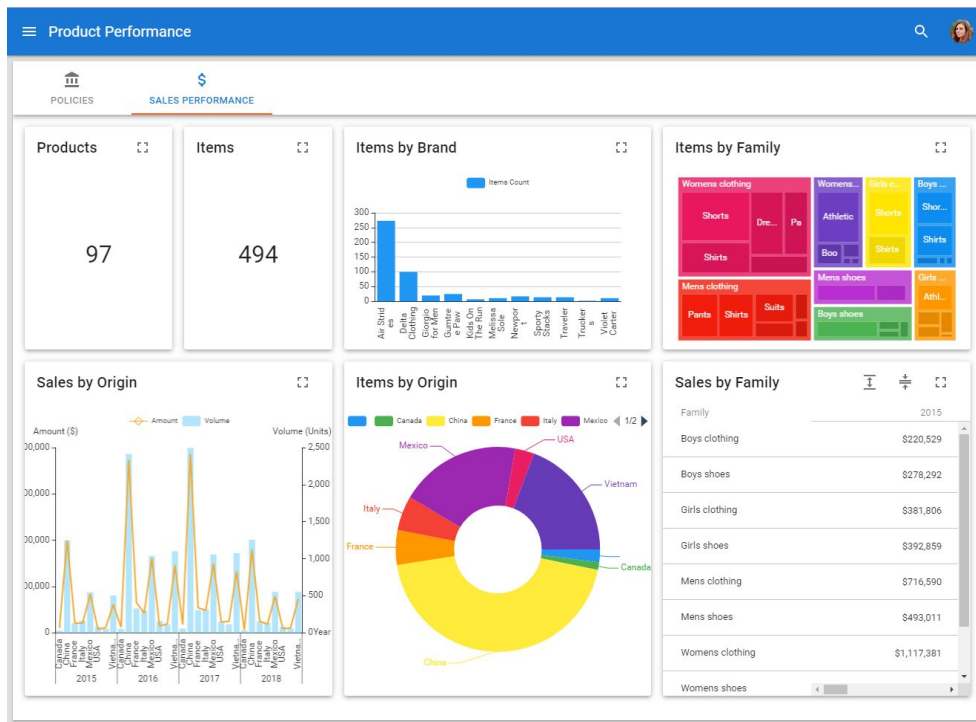
A sales performance dashboard blending data from the hub and from the data lake for sales figures

Information Governance and Policy Setting

xDM provides a design-time workbench interface into which IT and business users collaborate and design their data model, rules and policies. The workbench supports impact analysis and validation features to quickly react to governance decisions and apply updates to data rules, policies and stewardship functionalities.

Furthermore, xDM provides an extensible **Data Governance** application to manage business definitions and their actual instantiation as enforced rules in the Data Models.
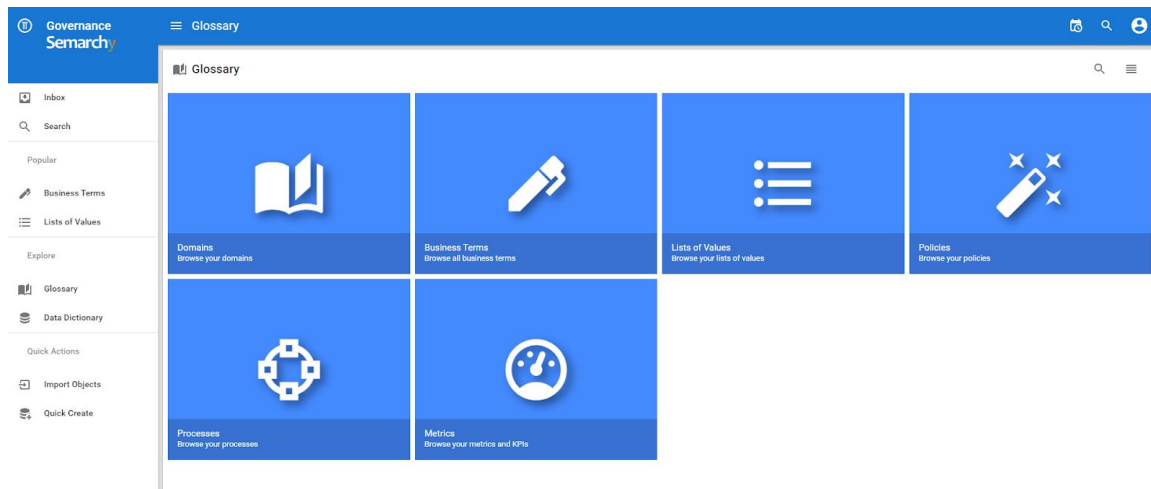
xDM Data Governance Application

The Governance Application includes a collaborative environment for business users and data champions to engage in end-to-end governance.

Enterprise Glossary

The application exposes an enterprise glossary to manage

- **Business terms** with their relationships, ownership, and RACI matrices;
- **Policies** and their relationships to data elements;
- **Processes, Tasks, Key Performance Metrics, Lists of Values** and their Mappings to existing Taxonomies;
- **Domains and contexts** to refine the glossary

Term definition and impact analysis in xDM Data Governance

Data Dictionary

The data dictionary allows capturing technical metadata from various endpoints to govern the physical deployment of applications. It manages information about:

- Servers, Database, Schemas
- Datasets, tables, files, views, dashboards
- Attributes, Columns, RI rules
- Relationships to the data hub models and glossary
- Ownership and RACI Matrices.

Users & Roles

The cornerstone of the Data Governance application resides in the definition of roles and users with their expected authorizations in the applications ecosystem. The governance application helps to document:

- Roles and authorizations;
- Access contexts, risks, and compliance policies;
- Business stakeholders and their roles in the governance processes (RACI)

Data Issues Escalation Workflows

The governance application includes customizable issue tracking processes leveraging the xDM workflow engine. Business users, while browsing, consuming, editing, or managing data can quickly escalate issues related to their contextual use of the data to the appropriate identified owners/subject matter experts.

Multiple implementation style support / Multiple domain support

xDM is designed from the ground up to be multi-vector (multi-style, multi-domain, multi-usage scenario, multi-organizations, cross-industries). As such, it natively supports the 4 implementation styles simultaneously within the same data model(s). Our customers have successfully implemented the 4 styles depending on their use cases using our platform.

Data models in xDM can be implemented in a single instance or in multiple instances with cross-domain intersections. All elements in the designed application (workflows, user experience, forms, collections, search patterns, roles, privileges, etc.) are uniform and interoperable.

The following table lists the relative usage of the hub styles for the most common domains implemented by our customers:
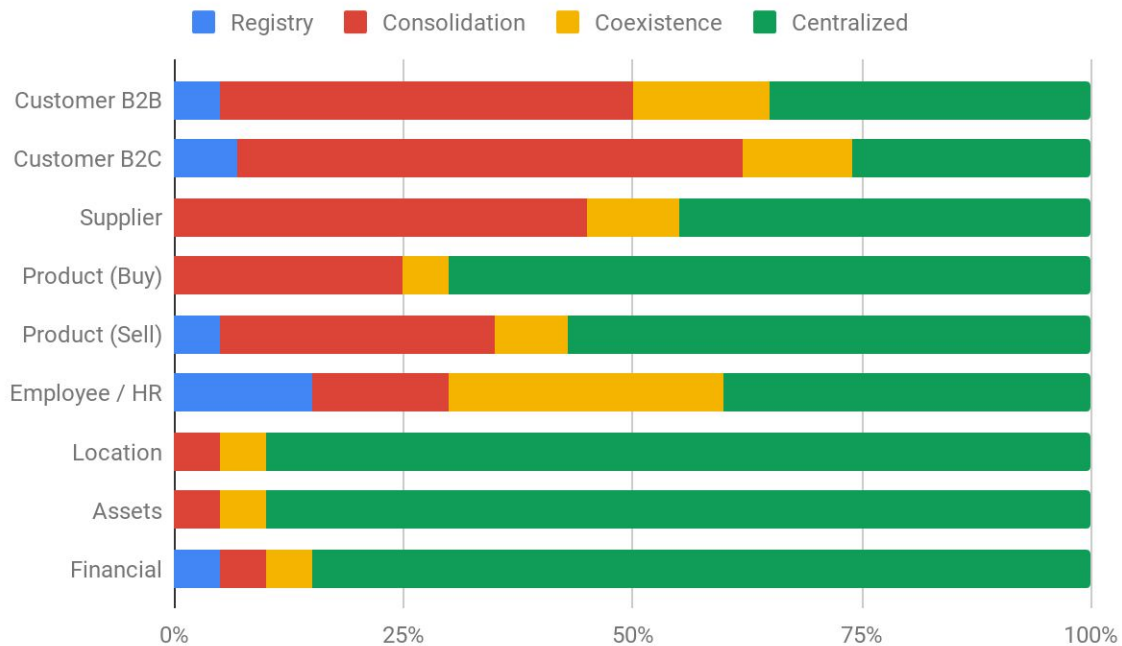
Implementation hub style per data domain

Refer to the Intelligent Data Hub™ section earlier in this document for details.

Multiple usage scenario (operational/analytical) support

xDM is primarily used by our clients to solve for operational usage scenarios. Most of these scenarios will also use xDM as the primary source for dimensional data to populate the analytical environment.

The Intelligent Data Hub™ becomes the central point for information lifecycle management and orchestrates business-critical processes for improving the overall data quality and semantics.

Most of our clients integrate the data hub with their CRM applications (Salesforce, SAP CRM, Siebel, HubSpot, SugarCRM, etc.), ERPs (SAP, INFOR, NetSuite, Oracle EBS, etc.), HR applications, Financial applications, eCommerce Applications, POS apps, etc. xDM tightly interacts with the enterprise processes (Order to Cash, Health and Environmental Compliance, Consent Management, Privacy and GDPR, IDMP, CFR 11, etc.) from within the calling application by invoking xDM services through REST APIs. For example, a shipping company using SAP decides to enhance the booking process to comply with export regulations with regards to party data. When a Sales Rep enters a booking in SAP, xDM captures the Party data and performs preliminary compliance checks. If the Party is determined as risky by xDM, the SAP transaction is suspended, and the Party data is submitted within xDM to the Regulatory Compliance teams for further screening. When the Party is finally approved in xDM, the Sales Rep is notified, and the booking transaction is resumed.

Similar use cases are implemented by our clients across most of the usual data domains.

Very few clients use xDM **only** for a purely analytical usage scenario (less than 10%). However, most clients using xDM in operational scenarios also leverage it for analytical use cases. Meaning, data from xDM is used as their source of trusted "golden record" data for Logical or Classical Data Warehouses as well as Data Lakes, predictive analytics and Big Data initiatives.

Data application suite - internal integration

By nature, xDM is a single and unified platform for designing data-driven applications starting from a data model and refining it down by adding processes, dashboards, profiling data, workflows, and user experience. Data Models share the same semantic layer and can eventually be deployed within the same single instance.

Our clients usually have more than one model for maintainability reasons and faster design cycles. They usually decide where to add the "intersection" data sets by selecting one instance as the primary management hub and the other instances as "slaves", where shared data is replicated or virtually accessed. Such replication or virtualization processes can be auto-generated by xDM Integrator or using any other ETL or ESB tool.

When clients decide for a given style for any of their base objects within a domain, they can easily switch to another style with minimal disruption. For example, a client might decide to design a centralized hub for Customer/Organization data where the Customer entity is defined as "Basic Entity" (see the Data Modeling section). During an M&A operation, and to streamline the integration of the Customer data owned by the second party, the client can decide to change the entity type to "fuzzy-matched" and load all the third-party Customer data into the hub. xDM would then naturally match and merge the 2 data sets and provide all data stewardship workflows. After this integration is performed, and when the client is ready to de-commission the Customer data applications of the acquired party, the client can switch back to the centralized mode again.

xDM Capabilities Detail

Product Solution Architecture

Semarchy xDM is an enterprise-scale integrated data hub unifying Master Data Management (MDM), Reference Data Management (RDM), Data Governance, Data Quality, and Data Integration.

The platform design is consistent across on-premises and cloud-based deployments.
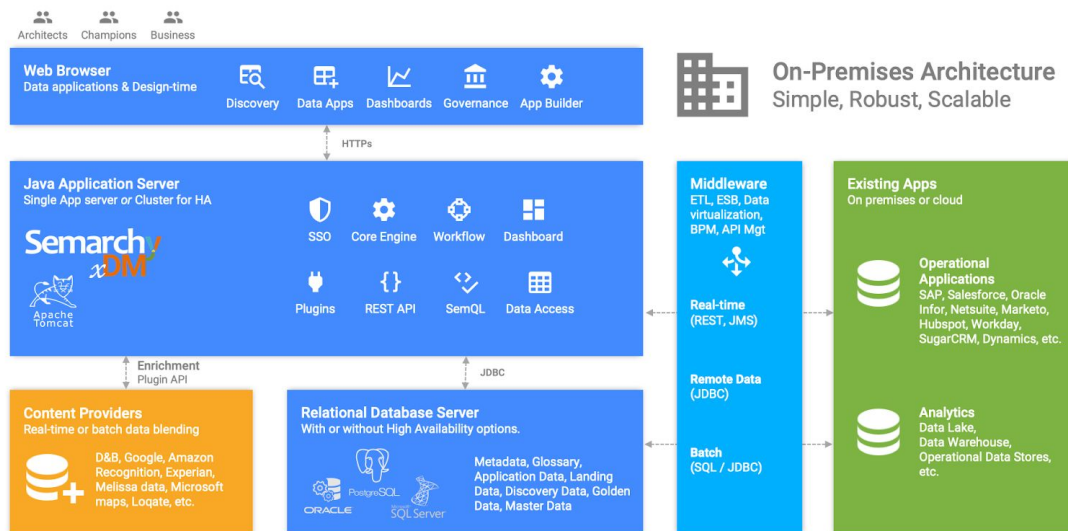
Reference architecture for on-premises deployment

The architecture of the xDM platform is a classical yet highly scalable three-tier architecture, using a Java EE application server with a relational database backend:
- User Interfaces, for design-time and business users, run in a **web browser**. They use the most recent front-end frameworks (Angular, React) for optimal user experience.
- These interfaces, as well as the REST APIs, are served by the **Java EE application**, which provides the interaction layer with the core MDM services while enforcing data and metadata access security.
- The **backend database** stores the xDM metadata (models, policies, etc.) as well as the hub data. This database is used to query the hub data and as the processing engine for the xDM Data Certification Process, storing the data along its lifecycle for full data lineage and traceability.

The platform exposes user interfaces, which all run in a web browser:

- Design-Time UIs enable application designers and data architects in their implementation of xDM data hubs:
  - o **xDM Discovery** to perform rapid source data profiling and analysis when starting the data hub initiative.
  - o **xDM Application Builder** to define the data models, information quality and policies, as well as the user interfaces and workflows for data stewardship and access.
  - o **xDM Dashboard Builder** to design charts, dashboards and applications that measure the quality of the data and efficiency of the processes managed in xDM data hub. Dashboards and charts can combine these metrics with metrics from operational or analytical data sources.
- Business users, Data Stewards and Data Champions UIs are automatically generated from the design and definitions created in the xDM Application Builder and xDM Dashboard Builder
  - o **xDM Data Management Applications** provide access to the data certified in the xDM data hub. It also provides management and stewardship capabilities, including workflows to author, curate, and maintain this data.
  - o **xDM Dashboard Applications** provide access to the dashboards and metrics for business users to measure and monitor the value of their data and of the data governance initiative. Dashboards and charts may also be embedded within the xDM Applications. Advanced users such as Data Champions and Data Stewards can also build their own dashboards depending on their privileges.
  - o **xDM Governance Application** is a customizable application for managing the enterprise glossary, policies, roles, issues, etc. (see Information Governance and Policy Setting)

In addition to the user interfaces, the following integration/programmatic interfaces are available out of the box.

- The **REST API** provides programmatic access to the capabilities of the xDM platform, including platform management, data integration and access, data authoring, and match and merge. Using this bi-directional API, integration specialists can plug their middleware on, or interact with, any moment in the data lifecycle, to access for example golden or master data, data quality violations or data history.
- Similarly, the **SQL API** provides access to the data at any point of its life cycle in the data hub and supports bi-directional master data exchange between xDM and any external system.

Plug-ins and **REST Clients** extend the capabilities of xDM to use third-party data sources and micro-services, enabled via RESTful APIs or direct integrations. Plug-ins and REST Clients are detailed in the [Enrichment / Standardization Rules](#) of this document. Clients and implementation partners can write their own customized plug-ins supporting business/industry specific logic or external data sources. Semarchy offers an open **Software Developer Kit (SDK)** that enables partners, clients, and any user to create bespoke extensions to our solution to fit their business needs.

xDM Integrator is available as an add-on to xDM. Its data integration (ELT) capabilities enable designers to easily define data mapping between any applications and xDM.

Development Environment

xDM provides a graphical approach to development, using a rich web-based interface for designing the data hub and applications, including data modeling, policy and information quality rules definition, workflows and user interface, API endpoints, actions, security, privileges, notifications, data subscription views, etc.

All design-time artifacts and metadata are stored in the xDM Repository. This repository supports multiple models (per functional domain) or a single model for all domains. This multi-user environment runs equally on-premises or in the cloud. Version control capabilities are built in the platform.
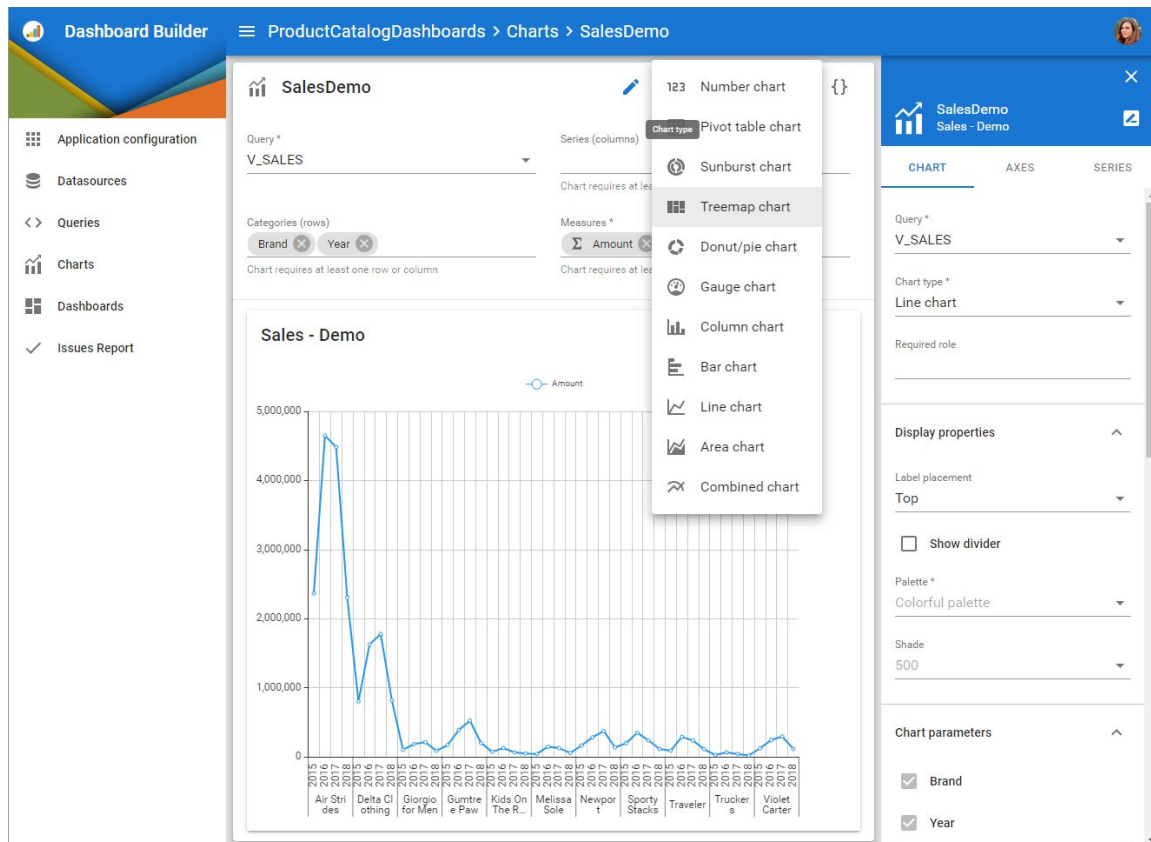
Example of the interactive design-time experience for building new charts in xDM Dashboards

Unicode Support, Language Support

xDM design-time UIs are available for developers in English and French languages.

The generated business user interfaces are internationalized for localization in any language using properties files.

xDM is designed and certified to support data in any language (incl. DBCS). Data is internally managed and stored in UTF-8 or UTF-16.

System Requirements

Application Server

Semarchy xDM Server is certified with Apache Tomcat 8.5.x and 9.0.x.
Semarchy xDM is a JEE7 web application and can also be installed in the following application servers:

- Eclipse Jetty 9.4
- IBM Websphere Liberty Profile 18.x, 19.x, 20.x
- Wildfly 17.x, 18.x, 19.x
- GlassFish 4.x, 5.x
- Oracle WebLogic 12c Release 2 (12.2.1.3.0)

More than 90% of our clients use Apache Tomcat.

RDBMS

The supported database versions for the repository and the data hubs are:

- **Oracle Database**: *Express, Standard or Enterprise edition version 12c, 18c and 19c, running on-premises or in Amazon RDS.*
- **PostgreSQL**: *version 9.6, 10, 11, 12, running on-premises, in Amazon RDS, Azure DB for PostgreSQL.* **Amazon Aurora** *versions 1.2 and above are certified.*
- **Microsoft SQL Server**: *Express, Standard or Enterprise edition version 2017 and 2019, running on-premises or in the cloud.* **Microsoft Azure SQL Managed Instance** *and* **Azure SQL Database** *are certified.*

xDM servers interact in real-time with the databases and intelligently balance between direct CRUD operations and Asynchronous writes with ACID compliance.

Data Storage and Legibility

Using xDM Integrator ELT capabilities or their own middleware/data integration layer, data architects can easily integrate xDM hub with external data sources, using direct SQL access to the data hub or the REST APIs.

The physical objects stored in the data hub schemas, as well as the REST API requests and response format are derived from the logical model's entities, attributes and relationships. Clients can access various documented data points at various stages of the data lifecycle. Refer to the [xDM Integration Guide](#) for further details.

Integration designers have access to a documentation automatically generated in HTML and Open API 3.0.
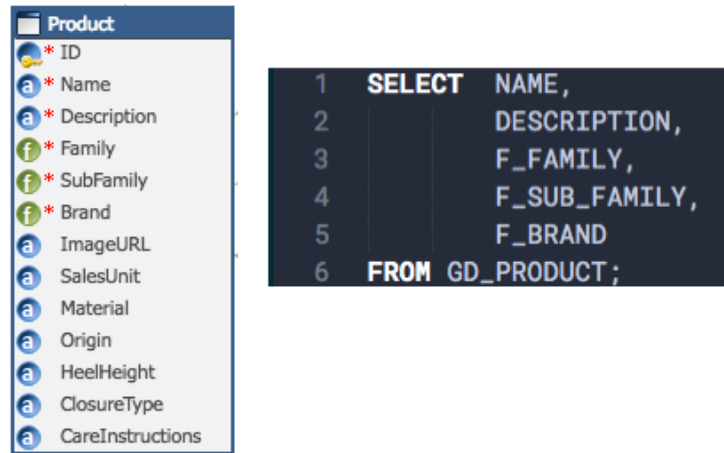
An xDM entity definition and the corresponding SQL query for Golden Data (GD)

Data Modeling, Mapping, Integration Patterns

Data modeling in xDM is typically performed on a client-case basis after a definition of the client requirements. The xDM development environment, due to its unique and single design-time UI and repository for all assets (governance, structures, policies, rules, workflows, UIs, hierarchies, etc.), easily supports an agile approach for the data management project, which typically runs in iterations.

The design exercise is often performed by a System Integrator or partner, although a substantial number of our clients use their own internal resources with the help of our Customer Success team to finalize their design iterations. Industry standard data models such as FHIR, FIB, ARTS, MS CDM, etc. can of course be freely leveraged.

Our Proof of Value service package allows our clients to use the pilot working application as the first iteration of their future data hub.

At the early stages of the project, most clients use the built-in import and export capabilities of xDM for inbound and outbound data flows. Once the data governance rules are well defined and business consensus is reached, the data integration team is involved to implement real-time or batch integration patterns via the inbound and outbound APIs (SQL, REST, JMS).

All our clients have implemented these integration patterns to support the various hub styles. Refer to Loading/Synchronization/Business Services/Integration for details.

Cloud-Based Deployment

xDM can be deployed on any cloud platform using the same code base and license.

Semarchy xDM is available as dedicated Microsoft Azure marketplace BYOL offerings: a standalone Quickstart VM to get started the solution, and a Solution Template for production deployments.

Example of an xDM deployment on Microsoft Azure

xDM cloud architecture uses the same concepts and code base as the on-premises product. Application servers are clustered with high availability (for example using Amazon EC2 instances or Azure VM Scale Sets) and the relational database store relies on the cloud vendor's platform capabilities (for example on Amazon: RDS for

PostgreSQL, Microsoft SQL Server, Oracle or Amazon Aurora). See System Requirements for more information about the supported Cloud databases.
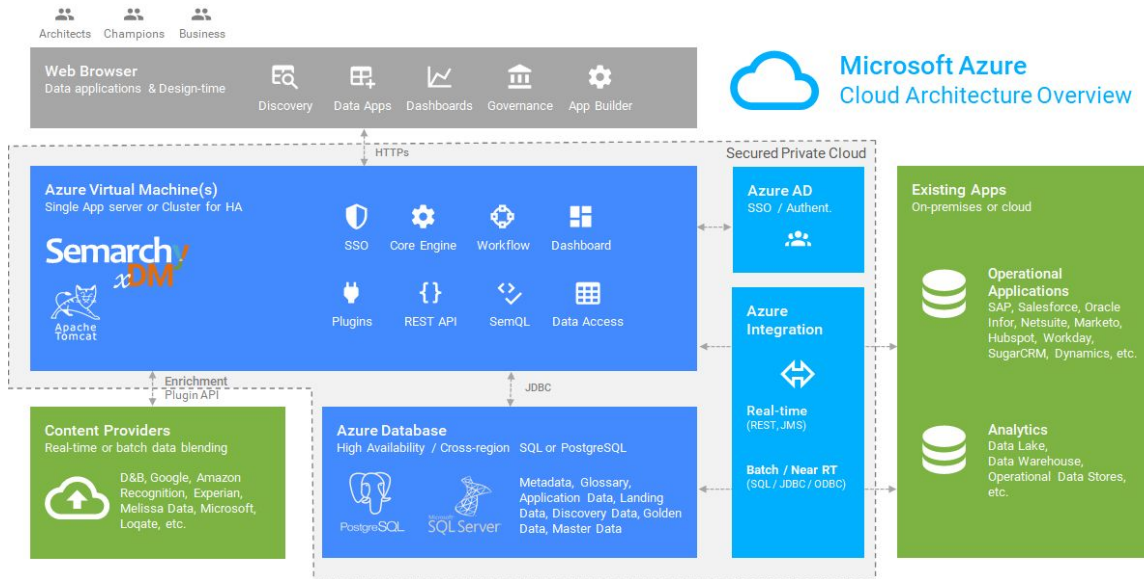
Reference architecture for cloud deployment (Azure example)

A dedicated Amazon Web Services setup is also available (EC2/RDS) with a BYOL (Bring Your Own License) model.

xDM is the most popular MDM and RDM solution on the AWS Marketplace for master data management.

Roadmap

Semarchy develops a single platform to help its clients manage their critical data assets in a value-driven approach.

As a high level, this approach involves the following steps:

- **Discover**: Discover and understand existing data to assess its quality and prepare for its governance.
- **Govern**: Define goals and metrics, refine glossaries, engage collaborators, assess compliance risks, and make decisions based on data discovery to iteratively build data-driven applications.
- **Integrate**: Define API endpoints, integrated authentication, external data hooks, and ensure integration performance is scalable, regardless of the protocol (REST, SQL, or JMS) used.
- **Manage**: Organize, integrate, store, enrich, validate, secure, and publish data that is consumed or produced by the enterprise-critical business processes. Have users collaboratively access and manage this data.
- **Measure**: Measure the value of the managed data and the efficiency of the processes in the context of the analytical and operational metrics.

Our strategic roadmap aims at delivering the components in the platform to execute these various steps while enhancing the designers' capabilities as well as data stewards and business users' productivity

The detailed roadmap for Semarchy xDM can be shared with our clients and partners under a non-disclosure agreement.

Support & Services

The Semarchy Proof of Value Approach

Semarchy has an intelligent go-to-market approach that we believe is differentiating in the MDM Market. When considering a master data project, organizations quickly face the challenge of evaluating ROI and value in order to gain executive sponsorship. Sponsors frequently perceive such projects as risky and time-consuming with low chances of success. They require convincing with solid facts.

The Proof of Value services package addresses this challenge.
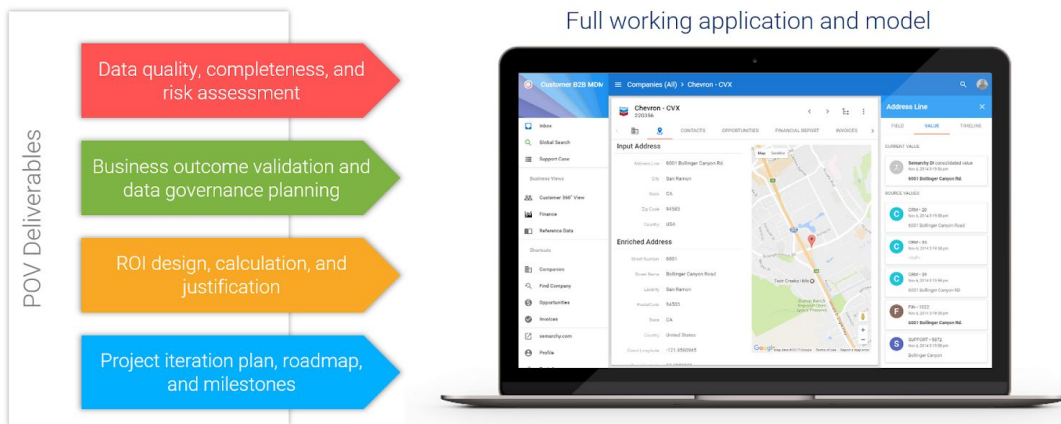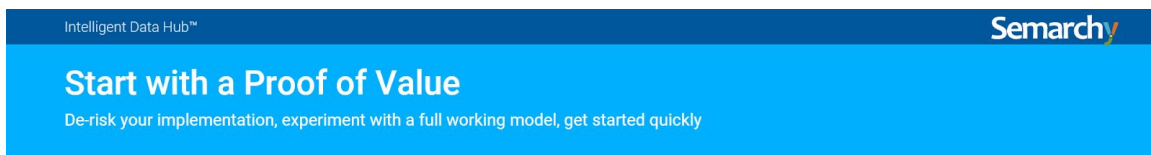
Semarchy Proof of Value deliverables

The Proof of Value process has enabled Semarchy to boast that 80% of our clients have a fully functional MDM solution up and running in 10 weeks or less. 60% of our clients have implemented more than one domain. 100% of our cloud-based subscription clients renew their solutions annually (or on three-year term agreements).

Customer Service and Support

Semarchy develops a community to help our clients and partners become self-sufficient using our platform. As such, Semarchy acts transparently with its technical materials and resources to have them publicly available. These include:

- [Downloads](#) or [Cloud quick start](#) of xDM
- [Comprehensive Online Tutorials](#) with detailed explanations and videos for data architects
- [Complete Product Documentation](#)
- [Community Forum](#) (public)
- [Online Service Desk](#) (restricted to clients) with over 200 Knowledge Base Articles
- Instructor-led classroom or virtual [training and certification services](#).

Product implementations by our clients and partners are monitored by our Customer Success Team that engages in the early stages of the projects to refine the business requirements and provide guidelines for the iterations.

On the contractual side, our standard technical support policy includes:

- Technical Assistance (EMEA and USA) continuously from 9:00 AM CET to 6:00 PM Pacific time on weekdays (18hrs out of 24) by phone and email, with immediate ticket assignment and a first technical answer within 8 hours.
- Support includes two levels of technical support plus escalation to the product engineering team.
- Custom contracts can also be designed per client requirements, such as 24/5 or 24/7 support and advanced SLAs.

Professional Services

Semarchy Customer Success Team leads our [professional services](#). This team engages very early in the implementation stages to mentor the client and/or partner resources. It is usually in charge of refining the business requirements, engaging in the early data governance workshops, helping design the first iterations, providing best practices, working with the business users to set up success criteria and key performance metrics, monitoring the completion and acting as the liaison between the clients and the Semarchy Product Management team.

Qualify

Customer Success attends **early** meetings to help Sales qualify business requirements

Showcase

Demo from Sales.

More POV or attractive custom Demos.

SHOWCASE our outstanding product.

Reuse

Customer Success, PM and Marketing teams builds library of **reusable assets** & accelerators.

Partner

Engage partners early when appropriate.

Work on technical **AND** commercials together.

Thanks to the agile nature of xDM, a typical project targeting a single domain with the integration of less than 5 applications would usually consume up to 10 man-days from the Customer Success team.

Semarchy primarily partners with delivery and implementation organizations with industry and domain-specialized focus/training (See Partners below).

For clients that decide to implement xDM using their internal resources, the Customer Success team provides tailored training based on data extracts provided by the client. This creative approach demonstrates the agility of the technology and allows the client resources to manipulate the platform in the context of their business requirements.

Partners

Semarchy boasts an international network of partners, who have been a key driver of growth, particularly over the past year. A list of key software, implementation, distribution, enrichment, catalog, and marketing service partners is available on our website.

Software Partners

Complementary software offerings that complete the data hub ecosystem are friendly partners for Semarchy. The nature of these relationships includes joint development conversations where appropriate, hooks/connectors or plug-ins, etc. Examples of such relationships include Denodo (Data Virtualization) and Stambia (Data Integration).

Implementation Partners

Semarchy is now engaged with over 60 implementation partners globally. Our focus is on data specialist partners with deep expertise, who engage with end-user clients in digital transformation, business intelligence, risk and compliance, and related projects. The company has also focused on extending partnerships and collaboration with some of the largest Global Systems Integrators.

Semarchy offers training, tutorials, free trial license software and helps new implementation partners get up and running quickly. Semarchy does not intend to grow an internal professional service delivery team, relying instead on partners to expand into new industry verticals, regions, and use cases.

Distribution/Selling partners

To extend the reach of primary offices in the US, UK and France, distribution partners (some of whom are also systems implementation partners) are located in Central and Latin America (Mexico, Brazil, Argentina, Chile), Asia (Japan, China, Australia, and Singapore), the Middle East, Africa, as well as Northern, Western and Central Europe (Benelux, Germany, Spain, etc.) to extend the reach of our offering with an indirect sales model.

Market Service Providers

Semarchy clients rely on our superior matching and merging capabilities, which can execute quality and enrichment tasks at enterprise scale. To facilitate this, several partners provide easily pluggable data feeds for processes such as address and phone verification. Examples of such partners include Melissa Global Intelligence and GB Group Loqate, Experian, Google, Dun & Bradstreet.

A full list and description of Semarchy Partners can be found at
www.semarchy.com/partners

About Semarchy

Semarchy is the **Intelligent Data Hub**™ company. Its xDM platform solves for enterprise Master Data Management (MDM), Application Data Management (ADM), and Collaborative Data Governance challenges at some of the most well-known brands across the US and Europe. This agile platform leverages smart algorithms and Material Design to simplify data discovery, stewardship, quality, enrichment, workflows, and measurements. xDM from Semarchy is implemented via an agile and iterative approach that delivers business value almost immediately and scales to meet enterprise complexity.

Semarchy xDM at a glance

Semarchy xDM™, Intelligent Data Hub™, Intelligent MDM™, Evolutionary MDM™, and Convergence™ are trademarks of Semarchy. Other names cited in this document are trademarks of their respective owners.

Disclaimer

This document is provided to you by Semarchy AS-IS and reflects the latest up-to-date information about our company and products. It is not a document endorsed by any analyst firm.