

# **Company & Product Overview**

Complimentary eBook Winter 2018

© Semarchy 2018

# Contents

Company	4
Overview	4
Intelligent Data Hub™	4
Smart	4
Agile	5
Measurable	5
Intelligent Data Hub strategy in practice	5
Business Operations	5
Intelligent Data Hub™	6
Overview	6
Intelligent Data Hub for Parties	7
Customer B2B, Legal Entities, Business Partners, and Suppliers	7
Customer B2C, Citizens, and Households	8
Employees, HR, Business Units	8
Intelligent Data Hub for Things	8
Product Data (buy/sell-side)	8
Financial Instruments, services, hierarchies, and securities	9
Assets	9
Intelligent Data Hub for Places	10
Intelligent Data Hub for Reference Data	10
Intelligent Data Hub for Data Governance, Data Discovery, and Data Catalog	11
xDM – Packaged Enterprise Software Product	12
Workflows and Business Processes	12
Key Capabilities	13
Loading/Synchronization/Business Services/Integration	14
Key Capabilities	15
Data Modeling	16
Key Capabilities	18
Information Quality and Semantics	18
Enrichment / Standardization Rules	19
Validation Rules	20
Match Rules	21
Survivorship (Merge) Rules	22
Performance/scalability, Availability and Security	23
Performance/scalability and Availability	23

	Security	24
	Hierarchy Management	24
	Information Stewardship and Policy Enforcement	25
	Information Stewardship	25
	Policy Enforcement	26
	Information Governance and Policy Setting	27
	Enterprise Glossary	27
	Data Dictionary	28
	Users & Roles	28
	Data Issues Escalation Workflows	
	Multiple implementation style support / Multiple domain support	28
	Multiple usage scenario (operational/analytical) support	29
	Data application suite – internal integration	29
xI	DM Capabilities Detail	31
	Product Solution Architecture	31
	Development Environment	32
	Unicode Support, Language Support	33
	Application Server Support	33
	OS/Hardware Server Support	33
	RDBMS Support	33
	Data Storage and Legibility	34
	Data Modeling, Mapping, Integration Patterns	34
	Cloud-Based Deployment	35
	Roadmap	35
S	upport & Services	37
	The Semarchy Proof of Value Approach	37
	Customer Service and Support	37
	Professional Services	37
Р	artners	39
	Software Partners	
	Implementation Partners	
	Distribution/Selling partners	
	Market Service Providers	

# Company

# Overview

Semarchy is the Intelligent Data Hub<sup>™</sup> 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 stewardship, quality, enrichment, and workflows. xDM from Semarchy is implemented via an agile and iterative approach that delivers business value almost immediately and scales to meet enterprise complexity. In fact, over 80% of Semarchy clients have a fully functional solution in under 10 weeks.

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 <u>https://www.semarchy.com/management-team</u>.

# Intelligent Data Hub™

Since Semarchy was founded in 2011, the software platform has expanded capabilities in a consistent manner, in line with the initial objectives of the business: to provide organizations with a single platform for multi-vector MDM, while solving for information stewardship/governance, data quality and enrichment, workflows, hierarchy management and reference data management in parallel.

As smart algorithms and semantic learning, as well as basic machine learning and Artificial Intelligence have been incorporated into the platform, we now refer to xDM as an "Intelligent Data Hub."

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. Semarchy believes that Data Management is about delivering enterprise-centric semantic consistency for any kind of data. Capabilities such as governance, data quality, enrichment and workflows are an integrated requirement, not one that should be cobbled together with different bits of software.

#### Agile



Agile methodologies applied to data applications are materialized via an intuitive, integrated design, development and user interface. Adaptive modeling enables business flexibility. Semarchy has taken the agile software development concept, along with the Material Design Language popularized by Google, and employed it for data hub applications.

#### Measurable



A data hub is as an initiative, not a technology. xDM implementations start with the business value before the technology. For this reason, xDM has been engineered so that it is evolved into a business with minimal disruption, to optimize adoption and show measurable enterprise ROI almost immediately.

#### Intelligent Data Hub strategy in practice



Application of Agile methodologies forces collaboration between all key stakeholders to define and deliver solutions with value very rapidly. Our <u>Proof-of-Value</u> (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.

### **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.

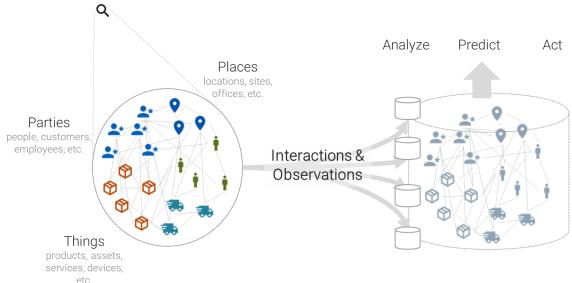
More information on the company mission, vision, and leadership can be found <u>on</u> <u>our website</u>.

# Intelligent Data Hub™

# Overview

The concept of an 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.).

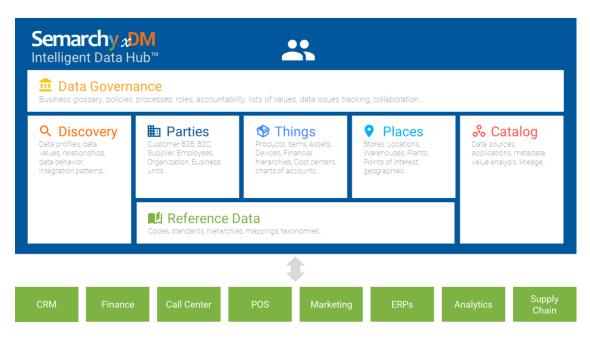
With the growing interactions and observations collected by enterprises through various data acquisition points, it is becoming critical to collaboratively govern these core data elements in an agile and measurable environment. Setting up this collaborative governance within an Intelligent Data Hub drives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



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.

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 <u>https://www.semarchy.com/clients</u>.

# 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, HIPPA, marketing segmentation, surveys, retention, loyalty, and predictive analytics processes.

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

#### **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 partially or fully be 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, 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 <u>Multiple implementation style support / Multiple domain support</u>.

#### **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 be partially or fully 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 the business requirements. They are detailed in the <u>Multiple implementation style support</u> / <u>Multiple domain support</u>.

#### 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, de-duplication, consolidation, collaborative information completeness, partner enrichment, classification, media enrichments, configuration, BOM, 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 the business requirements. They are detailed in the <u>Multiple implementation style support / Multiple domain support</u>.

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, de-duplication, 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 the business requirements. They are detailed in the <u>Multiple implementation style support / Multiple domain support</u>.

#### 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, debts, credits, risks, etc.

Critical business processes that can partially or fully be managed in the hub include asset creation, de-duplication, consolidation, classification, configuration, lifecycle and timelines 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 <u>Multiple implementation style support / Multiple domain support</u>.

# 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, de-duplication, 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 <u>Multiple implementation style support / Multiple domain support</u>.

### 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 echo system include taxonomy governance, code-to-code translation, global versus local versus industry definitions, industry standard on-boarding, 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



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

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
- Basic 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, **xDM**. Previous versions in the same code line were referred to as "Semarchy Convergence for MDM," after version 3.3, 4.x was renamed xDM.

Key features in the latest xDM release can be found here: blog.semarchy.com/xdm-4-4-features. Most notably, this version adds support of PostgreSQL.

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 "white board modeling." They are automatically deployed when the data model is instantiated. Workflows are version controlled with the model.

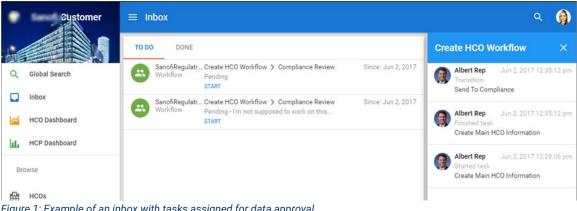


Figure 1: Example of an inbox with tasks assigned for data approval.

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 data hub.

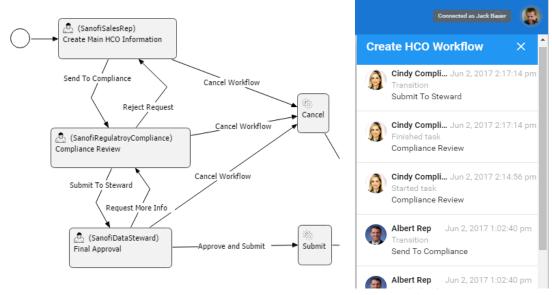


Figure 2: Multi-step approval workflow example (design-time) and traceability/collaboration at runtime

Once deployed, workflows are triggered on actions (such as Create, Edit, Mass Update, Delete, Import, etc.) and are available according to users' 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 reassigning their tasks.

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

#### **Key Capabilities**

- Policies/rules defined in the model are automatically enforced in the workflow
- Privileges for data access (at row level and at 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 time to complete. Non-compliance with SLAs can trigger special actions such escalation, reassignment, 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.
- 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 for building performance 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

*x*DM 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 end points called *named queries*. These then generate the appropriate APIs (REST or SQL), supporting complex structures as query parameters.

The <u>Semarchy xDM Integration Guide</u> gives further details about the supported integration patterns.



Figure 3: xDM Integration Pipeline

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 optimized for real-time access/low-volume per calls/high velocity through the *x*DM Continuous Loads asynchronous mechanism. The same process is also highly optimized for batch loading/high volume/low frequency type of integrations using indatabase set-based processing.

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

While many clients choose to use their own integration middleware, Semarchy offers an optional integration tool, *x*DM integrator, which works well for most usage scenarios.

ETLs and ESBs benefit from the SQL APIs as well as the REST APIs 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.

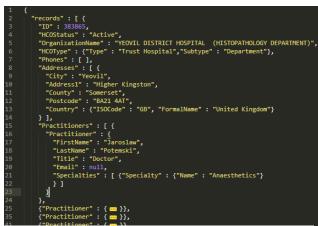


Figure 4: Example of a REST object generated by xDM

Querying REST APIs for reading or writing data respects the logical model semantics.

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

For example, using our REST API for massive real-time querying (30,000 requests per hour) on top of a medium sized data hub (12,000,000 golden records of customer data and about 45,000,000

source records) gives a constant response time of 25ms on average.

#### 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.
- 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 OEM Data Integration 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.

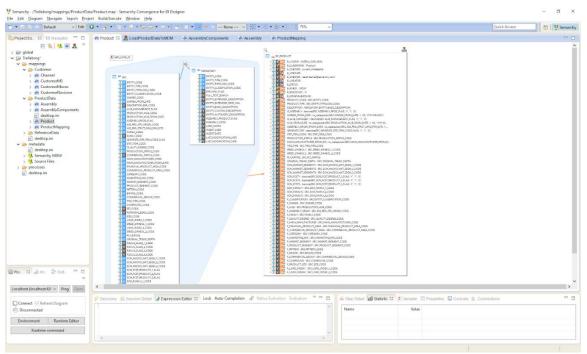


Figure 5: Example of a source to data hub Mapping in xDM Integrator

# **Data Modeling**

*x*DM 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.

*x*DM Data Models act as the single point of logical definition of entities, attributes, relationships, and rules managed in the data hub.

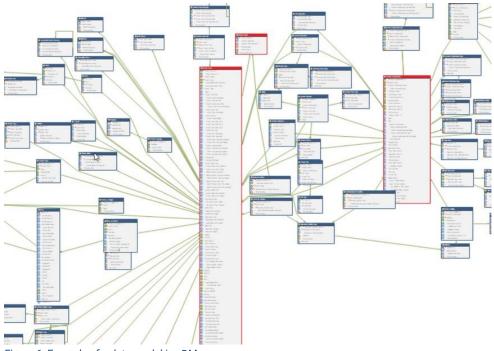


Figure 6: Example of a data model in xDM

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 *x*DM 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.

otio	Modify Entity	- x
Co	Entity	
Fu	Logical entity as defined by	the ER modeling standards
C		
in	Auto Fill:	<ul> <li>✓</li> </ul>
Co	Name:	Company
Co	Label:	Company
	Plural Label:	Companies
	Extends Entity:	<null></null>
	Physical Table Name:	COMPANY
rds:	Entity Type:	Fuzzy Matched 💌
ds:		Basic Fuzzy Matched
		ID Matched
,		
	Historize Golden Records:	✓
	Historize Master Records:	
		< Back Next > Finish Cancel

Figure 7: Definition of an entity in an xDM Data Model

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 bitemporal modeling (Historize Golden / Master Data) and automatically generate the appropriate effective dating and lineage data infrastructure and processes. **Key Capabilities** 

- Full multi-domain modeling framework supporting any end-user requirement.
- Support of reference data, master data and transactional/observational data modeling patterns.
- Support for all hub styles (registry, consolidation, coexistence and centralized)
- Support of graph-based modeling for complex relationship resolution. Refer to "<u>Does</u> <u>MDM Need Graph?</u>" blog for 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
  - o In-line data enrichment and data standardization rules.
  - ML-based rules for dynamic data classification.
  - External AI-driven micro-services 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**

The xDM <u>Data Certification Process</u> 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.

Performance of the data quality of the hub is monitored in the platform at any point in time. It can be exposed in dashboards within the *x*DM business user's application or using any external dashboarding/data visualization product.

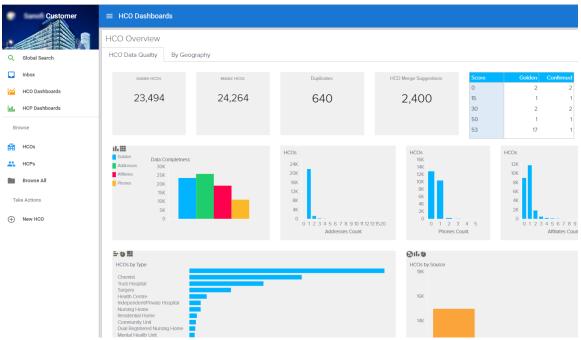


Figure 8: Example of a data quality dashboard in xDM

### **Enrichment / Standardization Rules**

The Data Certification Process uses the <u>data enrichment rules</u> to augment, cleanse or standardize the data using trusted external data providers or automated data transformations.

Enrichment rules can use Semarchy data manipulation language (<u>SemQL</u>) that allows for graph relationship resolution, data transformation and complex nodes filtering with a built-in library of 150+ functions.

					Can a sure s
SemQL Editor	r - Expression				<b>•</b> ×
		richer Expression NormalizedCompanyN that returns a value.	ame		
Attributes:	‡ª n+l ▼	SemQL Expression:		12 i	E 💛 🔌 👩 1xl (x)= 🍫 🚱
<ul> <li>Comp</li> <li>Norm</li> <li>Paren</li> <li>O Input</li> <li>Addre</li> <li>Rever</li> <li>Websi</li> <li>Norm</li> <li>Ticker</li> </ul>	anyName alized@ompan itCompany Address essHash nue iite alizedWebsite r	"(C#), '# if the name has 'Ba' ['('A), 'Islanks']; '('A 'find ''COMPORATION[CONF[C]]HOOS ['AA, '6](0, '9N, '8V, '8]('A) ['S, '8A, '8], '4]('A' 'Crech erfd ''S, '8A, '8](', '4] 'A' 'S point ''S, '8A, '8](', '4], '* Sponish ''S, '8A, '4](', '195, 'A' 'S and ''S, '8A, '4](', '195, '8')	<pre>ntity types. */ alizedCompanyNamed should already " before the entity type then DO any dot, comma, or spaces preced OUALEDINC[LATEDLITD[ASSOCAT] .~B0.*M1 *&amp; "CO[CL.*M1.*B1.*M1 *&amp; " * Fench entities */     /* French entities */ entities */</pre>	be ALL CAPS */ NOT strip off the entity Ing the business entity ON[ASSOC[L.vplL.vpl] *] *CO[L. ]*KG[* /* Ger	y type. For example: we wan type */ [l\.*t\.*t\.*P t\.*L\.*C P man entities */
<ol> <li>Sector</li> </ol>		Functions:	Variables:	Messages:	
	odedAddress	4 🛅 (All)	(x)= BusinessMatchThreshold	Description	Location
<ul> <li>Publis</li> <li>Source</li> </ul>		f(x) ABS	(x)= BusinessPublisherVariable		
<ul> <li>Source</li> <li>LoadI</li> </ul>		f(x) ACOS	(x)= BusinessValidationThreshold	=	
Classi		f(x) ADD_MONTHS	(x)= V_BATCHID		
0	ionDate	f(x) ADD_RECIPROCAL_RELATIONS	HIF (x)= V_LOADID		
-	teDate	f(x) ASCII f(x) ASCIISTR	(×)= V_USERNAME		

Figure 9: 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 calculation, value lookup, soundex and other metaphone algorithms, etc.

nricher: Modera	teImagesRekog	jnition 💣 🗸 🗶	
noFull 🔻 🛅 Product	▼ 🐉 ModerateImag	esRekognition 🔻	
<ul> <li>Name and Descri</li> </ul>	ption		
Name:	ModerateImagesRek	ognition	
Label:	Moderate Images Re	- kognition	
Description:		-	
Plug-in ID:	Amazon Rekonnition	Moderation Labels Enricher - com.semarchy.labs.amazo	n. 🔻
Plug-in Provider:	Semarchy		
Plug-in Description:	1	uses the Amazon Rekognition service to detect explicit a	n 🗆
Plug-in Legal Notice:	© Convright Semarch	by 2011-2017. All Rights Reserved. This program and the	a: (=
Enrichment Scope:	Pre-consolidation onl		
Position:	4	Ŷ	
Condition			
<ul> <li>Advanced Config</li> <li>Plug-in Params</li> </ul>	uration	G	8   3
Plug-in Inputs		fi	9   🕽
<ul> <li>Plug-in Outputs</li> </ul>		4	2   >
Attribute Name	nition.ExplicitNudity	Output Name ExplicitNudityScore	
-	nition.FemaleSwimw		
A ModerateRekog	nition.IsExplicitNudity	IsExplicitNudity	
A ModerateRekog	nition.IsFemaleSwim	IsFemaleSwimwearOrUnderwear	
🛷 ModerateRekog	nition.IsMaleSwimw	IsMaleSwimwearOrUnderwear	
🦽 ModerateRekog	nition.IsNudity	IsNudity	
🦽 ModerateRekog	nition.IsSuggestive	IsSuggestive	
Anderste Dekon	nition MaleSwimwea	MaleSwimwearOrUnderwearScore	-

Figure 10: Amazon Rekognition AI-based plugin

# **Validation Rules**

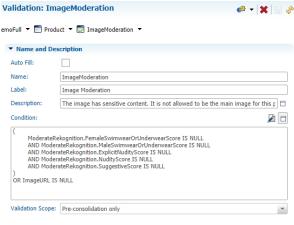


Figure 11: Data validation rule

*x*DM also provides a plugin framework to integrate with external data/micro-services providers. Those plugins fit in 2 categories:

**Content-lookup enrichment plugins**: B2B Party lookup (Experian, D&B), Address lookup (Google, Bing, <u>OpenStreetMap</u>), 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.

**Al/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.

Once the data is enriched by the Data Certification Process, it is validated against the <u>Data Validation Rules</u>. 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.

The 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 the business users or data champions for manual recovery.

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

· Product Authoring		
My New Product by Air Strides		
1 General Information		2 Image —
D		
16654		Description
Unique numeric product ID. Refer to internal guidelines.		
Name		
My New Product		
Use meaningful descriptions to optimize SEO		
Brand		
Air Strides (AIRSTRIDES)	$\times$	
Create a new Brand from the Brand menu		
Family		
Girls clothing (GIRLSCLOTHING)	×	
Validation failed for Subfamily must belong to selected family		
Sub Family		
Pants (Boys clothing)	×	

Figure 12: Example of a real-time data validation in the data authoring UI

#### Match Rules

	scription				
scr	ription: Company	name similarity			
Ma	tch Rules				x=ŷ 🚖 🐺
	Name	Label	Match	n On	Match Score
(=y	ExactNameAnd	Exact Name And Address	<cur< td=""><td>rent Entity&gt;</td><td>100</td></cur<>	rent Entity>	100
(=y	MatchEnriched	Match Enriched Address	<cur< td=""><td>rent Entity&gt;</td><td>90</td></cur<>	rent Entity>	90
(=y	FuzzyNameAnd	Fuzzy Name And Address	<cur< td=""><td>rent Entity&gt;</td><td>60</td></cur<>	rent Entity>	60
(=y	Website	Website	<cur< td=""><td>rent Entity&gt;</td><td>70</td></cur<>	rent Entity>	70
	rge Policy	from new master records:		75	
	e unconfirmed gold			75	
	-				
	e confirmed golder			75	
		confirmed golden records:		75	
		s to an unconfirmed golden r		75	
	new master record	s to a confirmed golden reco	rd:	75	
	ien master record.				

Figure 13: Configuration of multiple match rules with scoring

xDM provides a built-in <u>Fuzzy Matching</u> 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.

Multiple Match Rules in a matcher define several conditions for considering two records a match. Each condition has a different score that represents the percentage of confidence of the match.

Depending on the confidence score and the configuration of the merge policies, the matcher automatically merges groups of record together to create golden records.

When a match group is not merged automatically, because its confidence score is not high enough, it is flagged as a Merge Suggestion and queued for further screening by data champions. Merge Suggestions are reviewed by data

stewards with duplicate management actions, to decide whether to merge groups and create golden records.

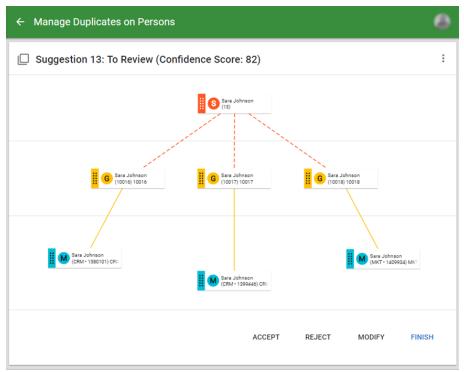


Figure 14: Example of the data steward UX for Merge Suggestions

#### 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).

Date of B	irth	×
FIELD	VALUE	TIMELINE
Name Label	DateOfBirth Date of Birt	
		nentation and to send
birthday cards		ffers
Group		th Rule (Date of Birth)
Override rule	Override un changes	til consolidated value
Consolidation rule		ublisher: (Marketing, Relationship nt, ERP)

<u>Survivorship Rules</u> defined in xDM by data architects and business users indicate which data survives in the golden records.

Such rules define how to consolidate the attribute values and the lifecycle of the user overrides when those values are edited/modified in the hub.

Consolidation rules propose 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.

Customer 02C Plus	<ul> <li>Manage Duplicates on Persons</li> </ul>				
FID DATE	Marquis Barayuga (10001) 10001	Marquis Barayuga (CRM + 1375249) CRM.1375249	Marquis Barayuga (CRM - 1388087) CRM.1388087	Markie Mark Barayug (CRM + 1987987) CRM.1987987	Marquis Bariyuga (ERP + 54354) ERP 54354
Global Search	Basic Information				
Inbox	10	10	10	10	iD
Start Here	10001	CRM.1375249	CRM.1388087	CRM.1987987	ERP.54354
wse Data	First Name	First Name	First Name	First Name	First Name
wse Data	Marquis	Marquis	Marquis	Markie Mark	Marquis
Customer	Last Name	Last Name	Last Name	Last Name	Last Nome
Products	Barayuga	Barayuga	Barayuga	Barayug	Bariyuga
Reference Data	Date of Birth	Date of Birth	Date of Birth	Date of Birth	Date of Birth
Browse Errors	Jan 30, 1976	Jan 30, 1976	May 10, 1994	Jan 31, 1976	Apr 4, 2017
	Person Type	Person Type	Person Type	Person Type	Person Type
ick Actions	Customer	Customer	Prospect	Prospect	Customer
Confirm Matches	Value Status	Value Status	Value Status	Value Status	Value Status
	HIGH	HIGH		NORMAL	HIGH
Review Suggested Matches	Name Enrichment				
Update Customers	These values are used for helping matc	These values are used for helping matc	These values are used for helping matc	These values are used for helping matc	These values are used for helping matc
Customer Count by Product					
annan a taran an an an 1978 (1979)	Phonetic First Name MRKS	Phonetic First Name MRKS	Phonetic First Name MRKS	Phonetic First Name MRKM	Phonetic First Name MRKS

Figure 15: Golden record with survivorship rules and value highlighting

# Performance/scalability, Availability and Security

#### Performance/scalability and Availability

*x*DM provides <u>an architecture for enterprise-scale deployments</u>, based on Java EE. *x*DM uses **PostgreSQL** 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 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. Semarchy clients include several very large-scale deployments.

	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 1-2 CPU/8GB RAM	Oracle Enterprise Edition PostgreSQL 2-4 CPUs/16-64GB RAM	Oracle RAC
Recommended RDBMS (cloud)	Amazon RDS db.m3.medium (Aurora, Oracle or PostgreSQL)	Amazon RDS db.m4.large (Aurora, Oracle or PostgreSQL)	Amazon RDS db.m4.4xlarge + (Aurora, Oracle or PostgreSQL)
Typical Application server configuration (per node) whether on-prem or cloud	Apache Tomcat 2GB RAM, 2CPU	Apache Tomcat 4GB RAM, 2CPU	Apache Tomcat 4GB RAM, 2CPU
Recommended Number of app server nodes (in the cluster)	2	2-6	6+

Typical configurations of our clients are summarized below:

*x*DM deployments are horizontally scalable by increasing the number of application server nodes in the cluster.

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

*x*DM 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.

*x*DM user interfaces and workflows 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 *x*DM 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 *x*DM organize entities into single or multiple related schemes for data management, classification, and reporting.

	Boissons > So	ft Drinks					م	8
Soft Drinks PGC · Boissons				<	>	E:	12	:
✓ 🕍 Nomenclatures SU	E:	=	÷					
V 🕅 PGC	CATÉGORIE	DÉTAIL	PRODUITS EAN					
🗸 🤹 Epicerie	Nomenclature SUID	111			Visuel			
🧐 Plats Preparés								
🔅 Pates alimentaires	Libelle	Soft Drinks					$- \Theta$	
V 🧧 Boissons	Statut	Actif					-1	
> 🝟 Soft Drinks	Catégorie parente	Boissons						
> 🔄 Alcools								
🗸 🌟 Surgelés, glaces								
> 🥒 Glaces								
> 🕟 Pizzas								
> 🥏 Plats cuisinés								
> 🧪 Légumes								
V 🎲 PFT								
🗸 🏀 Charcuterie & Traiteur								
🗸 鎭 Charcuterie								
> 🥜 Jambons								

Figure 16: 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 <u>SemQL</u> 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**

*x*DM provides a flexible and complete framework for generating a <u>Material Design</u> user experience for collaborating and managing the data of the hub. Applications generated by *x*DM 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 gate keeper of clean data.
- Expects metrics to measure performance of the processes and information guality.
- Spots and resolves data discrepancies with the help of the business stakeholders.



#### Data Champion

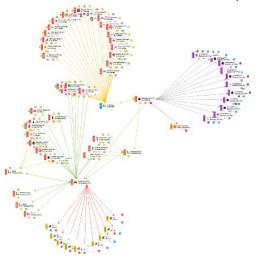
- Understands the impact of data changes across multiple business lines and applications.
- Performs data analysis using graph queries to reveal new relationships and opportunities.
- Understands and measures costs and risks related to bad data.
- Drives the data management roadmap whilst keeping a strong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



Figure 17: Example of a product 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.
- Manual review of match/merge decisions with possible 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.

Figure 18: Graph View 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 <u>Information Quality and</u> <u>Semantics</u>). 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.

# Information Governance and Policy Setting

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



Figure 19: xDM Data Governance Application

The Governance Application includes a collaborative environment for business users and data champions to engage in an 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

	usiness Terms > Invoid	ce			Sales and Marketing * Approved
ß	Invoice Sales and Marketing • Appro				
A none descri and de in cert and be	YOICE negotiable commercial instrumen bes, and quantifies the items sold livery and payment terms. ain cases (especially when it is si	t issued by a seller to a 8, shows the date of ship igned by the seller or sel paid in full. Types of inv	oment and mode of trans ller's agent), an invoice s roice include commercial	port, prices and discounts (if any)	
Relat	tes To (3)				
Relat	tes To (3) Object	Name	Description	Туре	
Relat	. ,	Name Purchase Order	Description A buyer- generated	Type Business Term	
telat P	Object Purchase Order		A buyer-		

Figure 20: Term Definition and Impact analysis

#### **Data Dictionary**

The data dictionary allows capturing technical metadata from various end points 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 echo system. The governance application helps documenting:

- 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 *x*DM 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

*x*DM 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 crossdomain intersections. All elements in the designed application (workflows, user experience, forms, collections, search patterns, roles, privileges, etc.) are uniform and interoperable.

Refer to Intelligent Data Hub<sup>™</sup> section earlier in this document for details.

# Multiple usage scenario (operational/analytical) support

*x*DM is primarily used by our clients to solve for operational usage scenarios. Most of these scenarios will also use *x*DM 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, SugarCRM, etc.), ERPs (SAP, INFOR, NetSuite, Oracle EBS, etc.), HR applications, Financial applications, eCommerce Applications, POS apps, etc. *x*DM gives the ability to tightly interact 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, *x*DM captures the Party data and performs preliminary compliance checks. If the Party is determined as risky by *x*DM, the SAP transaction is suspended, and the Party data is submitted within *x*DM to the Regulatory Compliance teams for further screening. When the Party is finally approved in *x*DM, 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. While few clients use *x*DM **only** for a purely analytical usage scenario, most use it as the source of trusted "golden record" data to populate Data Warehouses or Data Lakes directly sourcing from *x*DM.

### Data application suite - internal integration

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

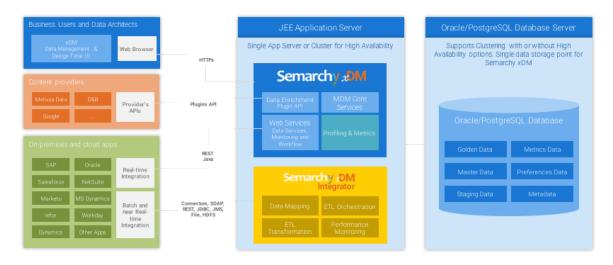
Our clients usually have more than one model for obvious maintainability reasons and designtime lifecycle. They usually decide where to add the "intersection" data sets by selecting one instance as the primary management hub and the other instances as "slaves" where intersection data is replicated or virtually accessed. Such replication or virtualization process can be autogenerated 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 <u>Data Modeling</u> 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" entity 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.



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 an 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 data hub 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:
  - 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.
  - **xDM Dashboard Builder** to design charts, dashboards and applications that measure the quality of the data and efficiency of the processes managed in *x*DM

data hub. Dashboards and charts can combine these metrics with metrics from operational or analytical data sources.

- Business users' UIs are automatically generated from the design and definitions created in the xDM Application Builder and xDM Dashboard Builder
  - **xDM Applications** provide access to the data certified in the xDM data hub, and to management and stewardship capabilities, including workflows, to author, curate and maintain this data.
  - 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.
  - xDM Governance Application is a customizable application for managing the enterprise glossary, policies, roles, issues, etc. (see <u>Information Governance and</u> <u>Policy Setting</u>)

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 *x*DM 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 lifecycle in the data hub and supports bi-directional master data exchange between *x*DM and any external system.

**Plug-ins** extend the capabilities of xDM to use third-party data sources and micro-services, enabled via RESTful APIs or direct integrations. Plug-ins are detailed in the Enrichment / <u>Standardization Rules</u> of this document.

Clients can write their own customized plug-in 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**

*x*DM 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 end points, actions, security, privileges, notifications, data subscription views, etc.

All design-time artifacts and metadata are stored in the *x*DM 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. And version control capabilities are built-in the platform.

# 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.

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

# Application Server Support

xDM Server is a JEE6 web application certified with the following application servers:

- Apache Tomcat 7.0.x, 8.0.x, 8.5.x and 9.x
- Eclipse Jetty 9.x
- IBM Websphere Liberty Profile 18.x
- Wildfly 8.x, 9.x, 10.x, 11.x, 12.x
- GlassFish 4.x, 5.x
- Oracle WebLogic 12c Release 1 (12.1.1, 12.1.2, 12.1.3), Oracle WebLogic 12c Release 2 (12.2.1.3.0)

More than 90% of our clients use Apache Tomcat.

# **OS/Hardware Server Support**

*x*DM server runs as an application in a supported application server. The hardware requirements are those of the application server.

Refer to Performance/scalability and Availability for other details.

#### **RDBMS Support**

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

- Oracle Database 11g Release 1: 11.1.0.6-11.1.0.7
- Oracle Database 11g Release 2: 11.2.0.1-11.2.0.4
- Oracle Database 12c Release 1: 12.1.0.1-12.1.0.2
- Oracle Database 12c Release 2: 12.2.0.1
- PostgreSQL version 9: 9.5.10-9.6.6
- PostgreSQL version 10: 10.1.x
- Amazon Aurora (PostgreSQL-Compatible)

*x*DM 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 <u>xDM Integration Guide</u> for further details.

Integration designers have access to a documentation automatically generated in HTML and Open API 3.0.



Figure 21: xDM Entity definition and 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.

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

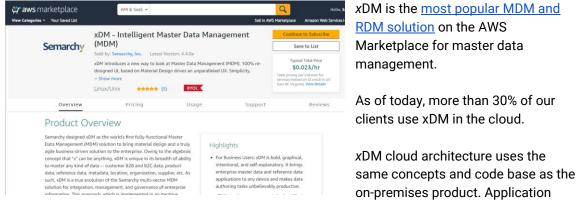
At the early stages of the project, most clients use the built-in import and export capabilities of *x*DM for inbound and outbound data flows. Once the data governance rules are well defined and business consensus is reached, the data integration team is involved to implement real-time or batch integration 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 <u>Loading/Synchronization/Business Services/Integration</u> for details.

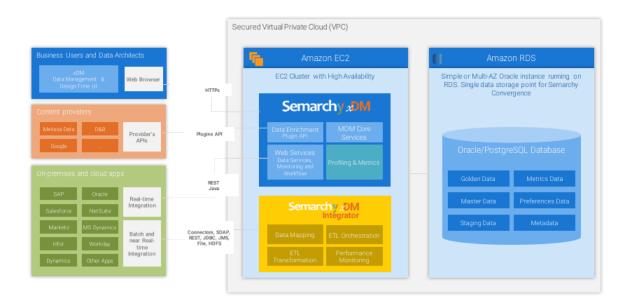
# **Cloud-Based Deployment**

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

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



servers are clustered with high availability (for example using Amazon EC2 instances) and the relational database store relies on the cloud vendor's platform capabilities (for example Amazon RDS for Aurora, PostgreSQL or Oracle).



# 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, setup metrics, refine glossaries, engage collaborators, assess compliance risks, and make decisions based on the data discovery to iteratively build data-driven applications.
- **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.

# **Support & Services**

# The Semarchy Proof of Value Approach

When considering a 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 <u>Proof of Value</u> services package addresses this challenge.

The Proof of Value process has enabled Semarchy to boast that 80% of our clients have a fully functional 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:

- <u>Downloads</u> or <u>Cloud quick start</u> of xDM
- <u>Comprehensive Online Tutorials</u> with detailed explanations and videos for data architects
- <u>Complete Product Documentation</u>
- Online Service Desk (restricted to clients) with over 200 Knowledge Base Articles
- In-class or virtual training services

The implementations of our product by our clients and partners are monitored by our Client 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 as stated in our contracts includes:

- Technical Assistance (EMEA and USA) 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 the 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 setup success criteria and key performance metrics, monitoring the completion and acting as the liaison between the clients and the Semarchy Product Management team.

Thanks to the agile nature of xDM, a typical project targeting a single domain with 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 *x*DM using their own internal resources, the Customer Success team provides a 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 18 months. A complete list of software, implementation, distribution, enrichment, catalog, and marketing service partners are available on <u>our website</u>.

# **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 engaged with over 20 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 also has relationships with three of the larger Systems Integration firms.

# **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.)

# **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 <u>Melissa Global Intelligence</u> and <u>GB Group Logate</u>, 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<sup>™</sup> 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 stewardship, quality, enrichment, and workflows. Organizations use the xDM platform to address challenges such as The Global Data Protection Regulation (GDPR). xDM from Semarchy is implemented via an agile and iterative approach that delivers business value almost immediately and scales to meet enterprise complexity. In fact, over 80% of Semarchy clients have a fully functional solution in under 10 weeks.

Semarchy xDM<sup>M</sup>, Intelligent Data Hub<sup>M</sup>, Intelligent MDM<sup>M</sup>, Evolutionary MDM<sup>M</sup> and Convergence<sup>M</sup> are trademarks of Semarchy.

#### Disclaimer

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