

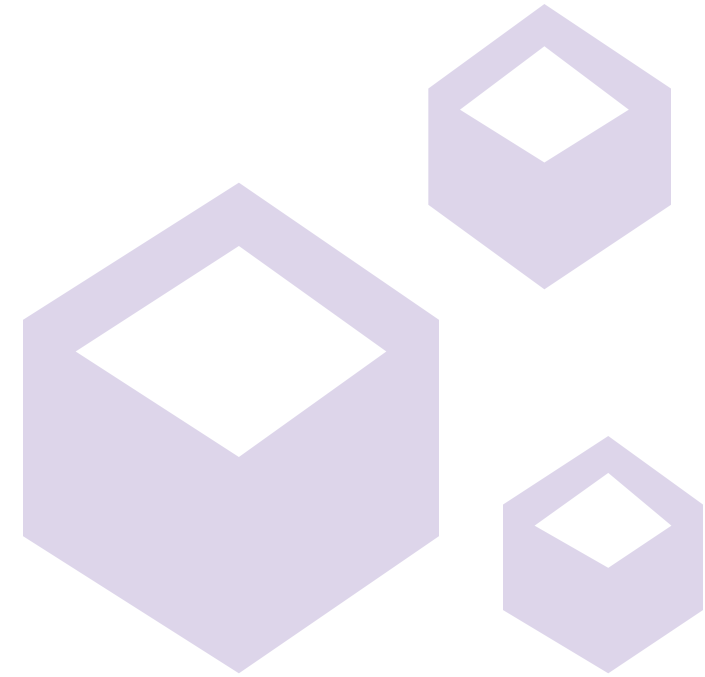
ataccama

# Journey to a Centralized Hub: Customer Cases Studies

MIT CDO IQ Symposium 2016  
July 13, 2016

# Agenda

- › Case Study 1: City Government MDM implementation
- › Case Study 2: Passenger Airline RDM implementation
- › Q&A





# Case Study

# CITY GOVERNMENT

# CITY | Project Highlights

## Primary drivers:

- › Data Sharing hub to serve all of the City's systems and departments (**Transactional MDM** shared entities – Party and Location – ensuring consistency and completeness)
- › **Tax / Revenue collection** and **Business register** are the main beneficiaries

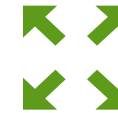
## Resulting in:

- › Single view of **legal entities** (both companies and entrepreneurs)
- › SOA architecture via **standard set of online services** providing master data

Solution live since 2013, currently being upgraded to make use of the **new Matching algorithms** and **stewardship interface** for **visual hierarchy browsing** editing the Master data entities.



Customer MDM and DQ



Operational / Consolidation style



Manual Editing  
Visual Hierarchy browsing



Online Integration / SOA



Identity Management Access Control

# CITY | Project Challenges

## Data



Data duplication



Ineffective Data Sharing



Lack of formal data governance



Application/function/business logic duplication

## Organizational



Complicated service structure



Complex organization

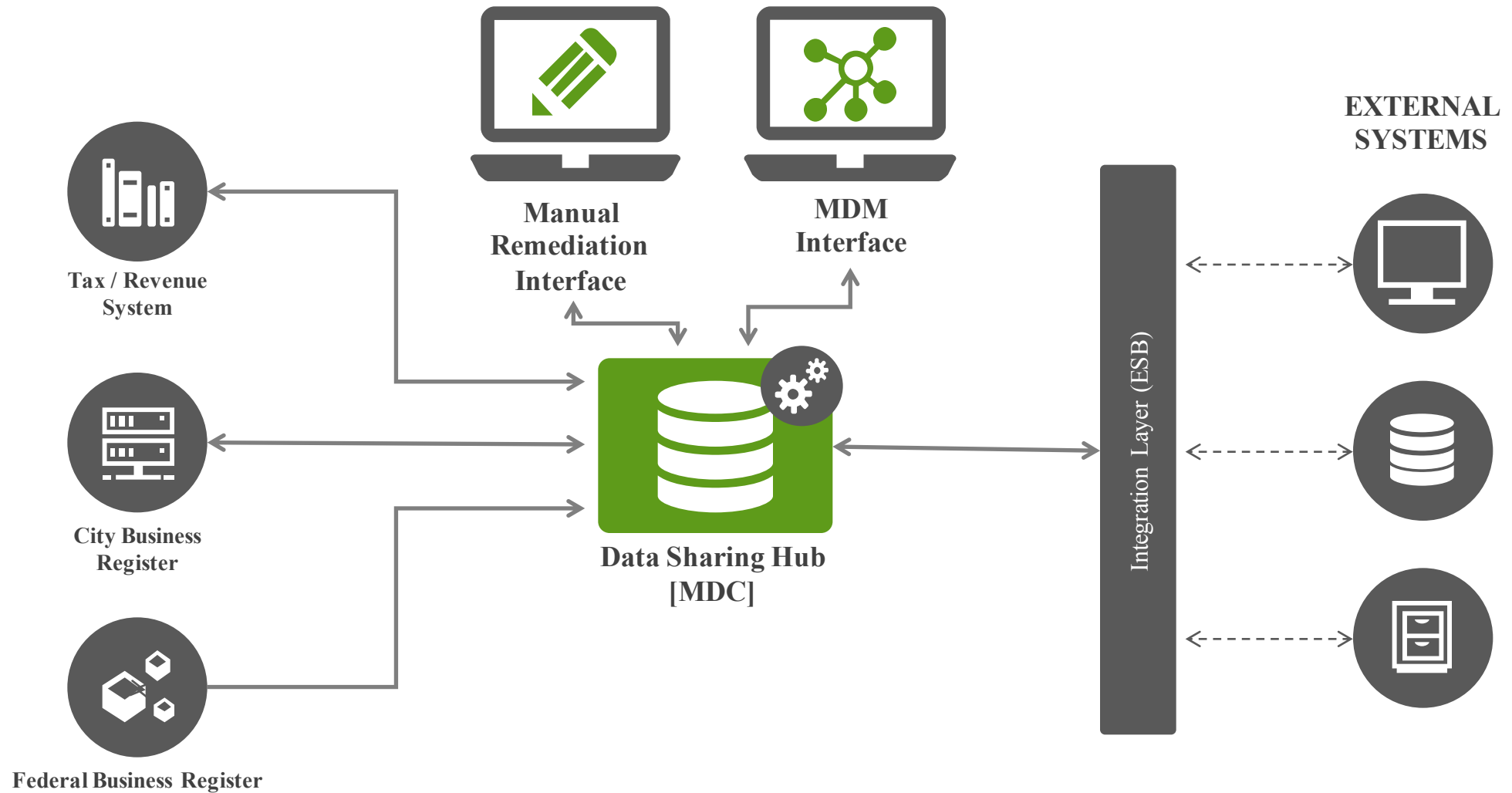


Corporate applications



Distributed IT governance

# CITY | High-Level Architecture



# MDM IDE

The screenshot displays the Ataccama MDC - masters Master Data Center 10.2.0 interface. The main workspace shows a logical model diagram with the following components:

- Entities (Yellow Boxes):**
  - contact**: id: long\_int, party\_id: long\_int, type: string, value: string
  - party**: id: long\_int, type: string, first\_name: string, last\_name: string, gender: string, birth\_date: day, sin: string, company\_name: string, business\_number: string, established\_date: day, legal\_form: string
  - address**: id: long\_int, party\_id: long\_int, type: integer, street: string, city: string, state: string, zip: string
  - id\_document**: id: long\_int, party\_id: long\_int, type: string, value: string
  - rel\_party\_party**: id: long\_int, parent\_id: long\_int, child\_id: long\_int, p2p\_rel\_type: string, rel\_category: string
- Relationships (Green Boxes):**
  - contact\_has\_instance**: connects **contact\_instance** to **contact**.
  - party\_has\_contact**: connects **party** to **contact**.
  - party\_has\_address**: connects **party** to **address**.
  - party\_has\_id\_document**: connects **party** to **id\_document**.
  - party\_has\_contract**: connects **party** to **rel\_party\_contract\_instance**.
  - party\_has\_parent\_party**: connects **party** to **rel\_party\_party**.
  - party\_has\_child\_party**: connects **party** to **rel\_party\_party**.
  - contract\_has\_party**: connects **rel\_party\_contract\_instance** to **contract\_instance**.
  - relation\_party\_has\_instance**: connects **rel\_party\_part...** to **rel\_party\_party**.
  - id\_document\_has\_instance**: connects **id\_document\_instance** to **id\_document**.
  - address\_has\_instance**: connects **address\_instance** to **address**.
- Instance Entities (Blue Boxes):**
  - contact\_instance**
  - party\_instance**
  - id\_document\_instance**
  - rel\_party\_contract\_instance**
  - contract\_instance**
  - address\_instance**
  - rel\_party\_part...**

The interface includes a menu bar (File, Edit, Navigate, Search, Project, Run, Window, Help), a toolbar, and several panels: Model Explorer, File Explorer, Instance Layer, and Outline. The Outline panel shows a hierarchical view of the model. The bottom status bar indicates "1 item selected" and "Licensed to: Tomas Vytiska".

# MDM Stewardship Interface

MDA Master Party Search in Master Party

Results for 'Jordan,ad' James Jordan | Master Party

**James Jordan | Master Party** high

James Jordan

**Basic Personal Information**

Party Id | 1,450

Customer Type | Person

First Name | James

Last Name | Jordan

Gender | Male

Birth Date | 1954 Nov 7

Social Insurance Number | 163679111

Age | 60

Active |

Updated | 2015 Sep 10 13:30:10

**Address**

Residential | 8

**ID Document**

ID Card | 7437

**Contact**

Email address |

Phone number |

Cell number |

**Contract**

Contract Type |

**Child Party**

Party | Caroline Jordan →

Party | Jane Smith →

**Instances: source columns**

Source System	Data source_id	First Name (src)	Middle Name (src)	Last Name (src)
<input type="checkbox"/> crm	101	James		Jordan
<input type="checkbox"/> life	301	J.		Jordan

MDA Product Search in Product

George

Frozen Sockeye Salmon Results for 'adv\*:active,r' AdventureWorks Cycles Results for '\*:active,not' Funds Investment | Prox Hierarchy | Product:287x

The diagram illustrates a network of entities and their relationships. Key entities include:

- PRUp [0/0] ID: 2873 Diamond Investment** (Product PRLo [2/2])
- PRUp [1/1] ID: 2875 Funds Investment** (Product PRLo [1/1])
- PRUp [0/0] ID: 2870 Progress Life Insurance** (Product PRLo [3/3])
- PRUp [3/3] ID: 2874 Invest Investment** (Product PRLo [0/0])
- PRUp [2/2] ID: 2872 Classic Insurance** (Product PRLo [0/1])
- PRUp [1/1] ID: 2871 Harmony General Insurance** (Product PRLo [1/1])
- Instance Contract ID: 2238 LINS** (Pry [1/1])
- Instance Contract ID: 2239 LINS** (Pry [2/2])
- Instance Contract ID: 2245 LINS** (Pry [1/1])
- Instance Contract ID: 2244 LINS** (Pry [1/1])
- Master Party PUo [0/0] ID: 2628 John Smith** (Adr [0/0])
- Master Party PUo [1/1] ID: 2629 James Jordan** (Adr [1/1])
- Master Party PUo [0/0] ID: 2645 Jane Jordan** (Adr [1/1])
- Master Party PUo [1/1] ID: 2643 Gery Boland** (Adr [1/1])
- Master Party PUo [0/0] ID: 2630 Caroline Jordan** (Adr [1/1])

Relationships shown include 'package', 'Risk', 'Capital', 'owner', and 'Home'.

Ataccama



# Manual Resolution Interface

## Customer data

### Original data

Processing Date: 2012-01-31 00:00:00.000  
Instance: HBFBIH  
Customer ID: 310796818739231  
Country: BA  
Segment B2: PK  
Segment Market: RM  
Active: N  
Obligo:  
On Balance Assets:  
# of Duplicates: 0

Severity	Message
Low	B2O_Obligo is missing (null).
Low	B2O_OnBalAssets is missing (null).

### Corrected data

Processing Date: 2012-01-31  
Instance: HBFBIH  
Customer ID: 310796818739231  
Country: BA  
Segment B2: PK  
Segment Market: RM  
Active: N  
Obligo:  
On Balance Assets:

Severity	Message
	No messages reported.

Comment:

- › Manage and resolve **DQ exceptions**
  - › Cleansing
  - › Match & Merge
- › **Drill-down** to source level
- › Complete **workflow** for data stewards
- › **DQ Firewall** to prevent new errors
- › **Progress reports** for managers

Filter

Severity:

System:

Entity:

Score:

Between  and

Number of records:

Between  and

Assignee:

Author:

Customer Status:

Record ID:

Between  and

Last name:

**For resolution**

[Actions: Bulk assign](#) | [Bulk transition](#)

Id	Sev.	System	Entity	Score	#Rec.	Data	Description	Assignee	Changed
101	✖	Sales	Customer	1510	1	Peter Taylor	Cleansing: Multiple issues	Alice	2012-07-27 17:00:12
102	✖	Sales	Customer	1160	1	Janet Holmberg	Cleansing: Multiple issues	Alice	2012-07-27 17:00:12
104	✖	Sales	Customer	1060	1	J. Holmberg	Cleansing: Multiple issues	Alice	2012-07-27 17:00:12
105	ℹ	Sales	Customer	60	1	Watson Jane	Cleansing: Multiple issues	Alice	2012-07-27 17:00:12
106	⚠	Sales	Customer	100	3	Regina Winters	Manual telephone selection	Alice	2012-07-27 17:01:13
108	⚠	Sales	Customer	100	3	Regina Winters	Manual selection	Carol	2012-07-27 17:01:13
109	⚠	Sales	Customer	100	2	Jane Black	Manual selection	Carol	2012-07-27 17:01:13
110	ℹ	Sales	Contact	0	1	Jane Black	Manual email verification	Bob	2012-07-27 17:01:13
111	ℹ	Sales	Contact	0	1	Regina Winters	Manual email verification	Bob	2012-07-27 17:01:13
112	ℹ	Sales	Contact	0	1	Jane Holmberg	Manual email verification	Bob	2012-07-27 17:01:13
113	ℹ	Sales	Contact	0	1	Wesley Nives	Manual email verification	Bob	2012-07-27 17:01:13
114	⚠								
115	⚠								
116	✖								
117	✖								

### Manual telephone selection

Record:	ID: 711	ID: 718	ID: 712	Preview
System:	Sales	Sales	Sales	
First name:	Regina	Regina	Regina	Regina
Last name:	Winters	Winters	Winters	Winters
Title:				
Gender:	F	F	F	F
Birth date:	1982-05-12	1982-12-05		1982-12-05
Social Insurance Number:	584279244	584279244	584279244	584279244
Telephone:	(203) 918-1162	(203) 488-8648	(203) -	(203) 488-8648

# CITY | Solution Benefits



Central data model



Easy management of MDM repository



Application integrations



Manual issue resolution



Automatic cleansing and consolidation



Compatibility with legacy systems



Service-oriented architecture



Scalability



**Case Study**  
**PASSENGER**  
**AIRLINE**

# AIRLINE | Project Highlights

## Primary drivers:

- › The EDW's mission is to provide **integrated business information**
- › **Reference Data Management** is identified as critical factor to the EDW's success
- › Simplify Code Maintenance support to alleviate load on EDW cycle
- › Improve Business buy-in by allowing direct interaction with the solution

## Resulting in:

- › Single consolidated hub for **enterprise reference data**.
- › Reduced IT costs
- › Faster response to changes and additions to reference data

Solution live since 2014, currently being upgraded to make use of Online Web-services. **In production with over 40 entities.**



Centralized Reference Data Repository



Back Propagation to Original Source Systems



Data Governance Centric Implementation



Built-in Data Quality



Identity Management Access Control

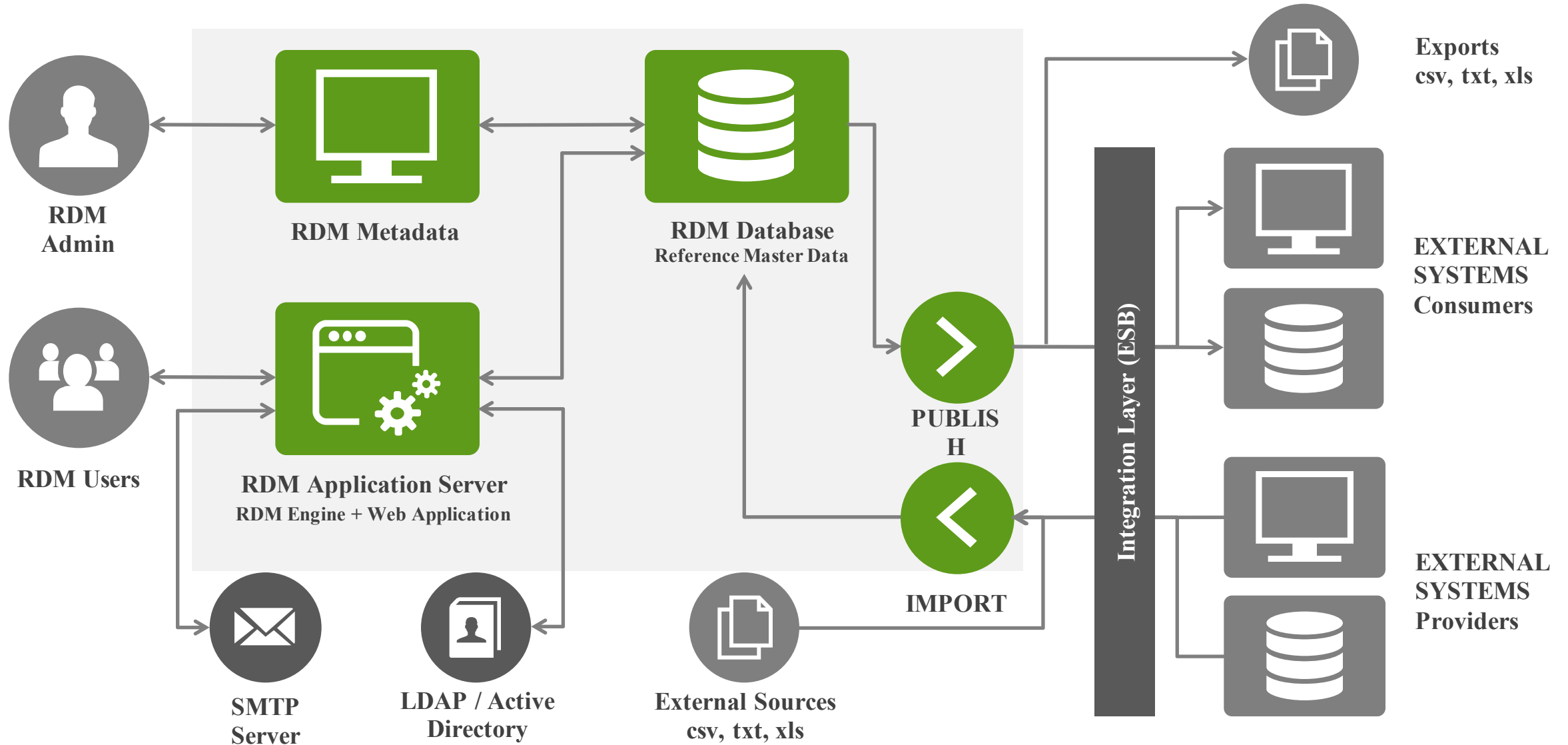
# Challenges of User Managed Reference Data

- › Manually maintained reference data drive company analytics
- › Lack of standard change management process
- › Business rules/validations are not supported, leading to faulty published data
- › No audit trail, review or approval
- › Sometimes ownership is not clear

# Challenges of User Managed Reference Data

- › Difficult to automate
  - › Difficult for systems to access/consume
  - › Difficult to implement a security policy
  - › Not all systems are consistently informed of the latest changes, resulting in discrepancies
- › Stored in multiple places
  - › Leads to duplication
  - › Introduces inconsistency
  - › Difficult to manage reference data at the enterprise level
  - › No process for agreeing on changes to reference data across the enterprise

# AIRLINE | High-Level Architecture





# Flight Range Reference Data Before

- › Data Steward manually maintained flight range data via spreadsheet
- › Not formally versioned
- › Distributed via email
- › Everyone ended up with their own copy
- › Did not handle merger

# Flight Range Reference Data Now

- › Easy to see who is maintaining the data
- › Unpublished records easily identified
- › Invalid records are easy to spot

Flight Range

Mode: EDIT

<input type="checkbox"/>	Flight Ran	Flight Rang	Effective	Expiration	Flight Range Description	Username	State	Valid
<input type="checkbox"/>	1	399	2014 Apr 16 0	9999 Dec 31 (	AA Mainline	500994	Edited	valid
<input type="checkbox"/>	900	1,699	2014 Apr 15 00:0	9999 Dec 31 (	AA Mainline			valid
<input type="checkbox"/>	1,700	2,099	2014 Apr 15 00:0	9999 Dec 31 (	AA Codeshare on US Mainline (			valid
<input type="checkbox"/>	2,200	2,499	2014 Apr 15 00:0	9999 Dec 31 (	AA Mainline			valid

# Flight Range Reference Data Now

**Edit detail**

Flight Range Number Begin	1	*	valid
Flight Range Number End	399	*	valid
Effective	2014 Apr 16 00:00:00	*	valid
Expiration	9999 Dec 31 00:00:00	*	valid
Flight Range Description	AA Mainline	*	valid
Operating Airline Code	AAz		INVALID
Flight Range Type	A	* valid	
Flight Range Category	PS	* valid	

Buttons: Enrich, Validate

All messages	Operating Airline Code	Global messages
<b>Validated</b>	<b>Name</b>	<b>Message</b>
INVALID	Operating Airline Code	String is 3 long but it can't be longer than 2.

**Lookup from 'Flight Range Category'**

Filter Columns setup

Flight Range Category	Flight Range Category De
PS	Passenger Scheduled Service
ES	Extra Section
CH	Charter
NR	Non-Revenue

1-4 of 4

Buttons: Set to null, Select, Cancel

› Instant validation

› Pick list support

# AIRLINE | Today and Beyond

## **In Production with over 40 entities**

- › Flight Range Codes
- › Flight Delays Codes
- › Seating Charts
- › Cargo
- › Flight Crew
- › Operations Planning
- › Operations Research
- › Revenue Management
- › Sales
- › Tech Ops

## **Current Development**

- › Customer Planning
- › Merchandising

## **Future Product Feature Adoptions**

- › Web Services
- › Operational Support Automation

# AIRLINE | Features & Benefits

## **Data Governance**

- › Standard maintenance process and workflow
- › Support auditing and data ‘versioning’
- › Improved data quality (business rules enforced before the data is published)

## **UI Designed with Data Steward in Mind**

- › Allows the business to directly manage their data!
- › Role-based UIs for review, approval, and publishing of changes

## **The Business and IT are Both Happy!**

- › Reduced IT costs
- › Faster response to changes and additions to reference data



# Ataccama Integrated Platform

# ATACCAMA INTEGRATED PLATFORM

## DATA GOVERNANCE PORTAL | Web Applications

Reference Data  
Manager  
(RDM)



DQ Issue  
Tracker  
(DQIT)



DQ Dashboard  
(DQD)



Master Data  
Admin  
(MDA)



Administration  
Web Console



## DATA MANAGEMENT | Engines

Master Data Center (MDC)



Data Quality Center (DQC)



DQ Analyzer (DQA)



Big Data Engine  
(BDE)



## SHARED PLATFORM COMPONENTS

Common  
Metadata  
Layer



Batch  
Interfaces



Online  
Interfaces



Adapters



Metadata  
Editor

Workflow



High  
Availability



# CITY GOVERNMENT PRODUCT STACK

## DATA GOVERNANCE PORTAL | Web Applications

Reference Data  
Manager  
(RDM)

DQ Issue  
Tracker  
(DQIT)

DQ Dashboard  
(DQD)

Master Data  
Admin  
(MDA)



Administration  
Web Console



## DATA MANAGEMENT | Engines

Master Data Center (MDC)



Data Quality Center (DQC)



DQ Analyzer (DQA)



Big Data Engine  
(BDE)



## SHARED PLATFORM COMPONENTS

Common  
Metadata  
Layer



Batch  
Interfaces



Online  
Interfaces



Adapters



Metadata  
Editor

Workflow



High  
Availability





# AIRLINE PRODUCT STACK

DATA GOVERNANCE PORTAL | Web Applications

Reference Data  
Manager  
(RDM)



DQ Issue  
Tracker  
(DQIT)

DQ Dashboard  
(DQD)

Master Data  
Admin  
(MDA)



Administration  
Web Console



DATA MANAGEMENT | Engines

Master Data Center (MDC)



Data Quality Center (DQC)



DQ Analyzer (DQA)



Big Data Engine  
(BDE)



SHARED PLATFORM COMPONENTS

Common  
Metadata  
Layer



Batch  
Interfaces



Online  
Interfaces



Adapters



Metadata  
Editor

Workflow



High  
Availability





# Key Differentiators

# Key Differentiators

---

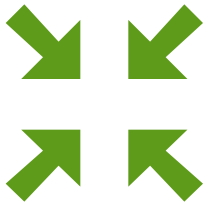
## EASE OF DEPLOYMENT



Business Centric UI



Data Model Driven



Truly Integrated  
Platform

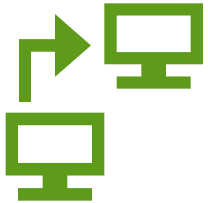


Reusable Components

# Key Differentiators

---

## STRONG TECHNOLOGY



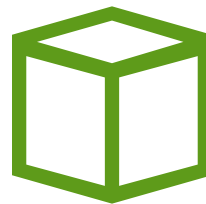
Service Oriented  
Architecture



Easy Installation  
and Maintenance



Performance



“White Box” Approach