



MDM Enterprise Analyzer: a framework to support centralized and local master data quality analysis

MIT IQ Industry Symposium,
Cambridge, Massachusetts, USA

by Kai-Uwe Baryga

SYDECON
Systems Design & Construction GmbH
Hans-Urmiller-Ring 46
D-82515 Wolfratshausen
Germany
Email: kai-uwe.baryga@sydecon.de

www.sydecon.de

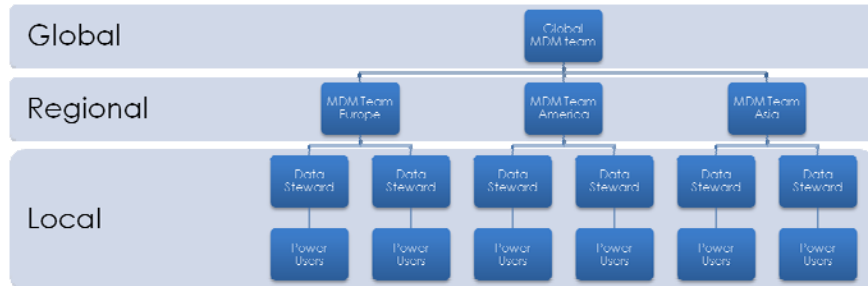


Master Data Key Success Factors Strategy and Governance

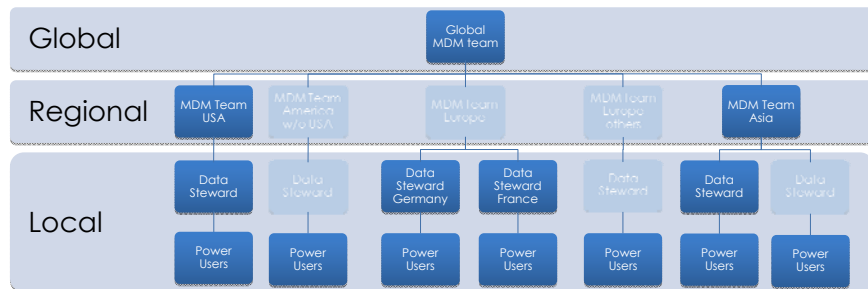
- Global MDM Management
 - Master data strategy and organization
 - Global master data quality and service controlling
- Global MDM Teams
 - Definition of global standards for master data objects and processes
 - Definition of global maintenance rules
 - Definition of global quality and service standards
 - Collaboration with regional MDM teams and global business drivers
- Regional MDM Teams
 - Definition of regional standards for master data objects and processes
 - Definition of regional rules in addition to global rules
 - Regional quality and service controlling
- Data Stewards
 - Collaboration with local business and MDM power users
- MDM Power Users
 - Maintenance of certain MDM objects or object parts



Hierarchical MDM Organizations



Typical MDM Organizations



- Incomplete, unbalanced hierarchies
- Inhomogeneous MDM experience and skills



Why improve MD Quality?

- Transactional processes base on MD objects, i.e. insufficient MD quality leads to ...
 - ... delays in the supply chain
 - ... expensive additional manual work
 - ... increased risks (credit limit, dangerous goods)
 - ... issues with reporting and analysis systems
- Demands due to
 - ... local laws and legal restraints
 - ... customer or vendor requirements
 - ... internal company policies
 - ... restraints required due to stock exchange (e.g. NYSE/SEC)

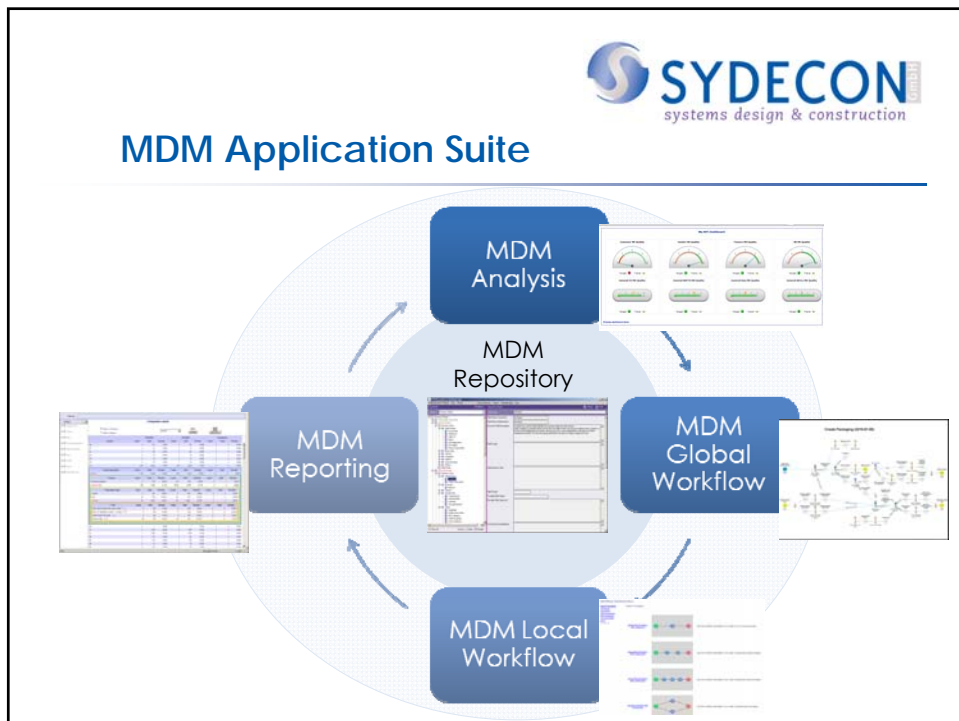
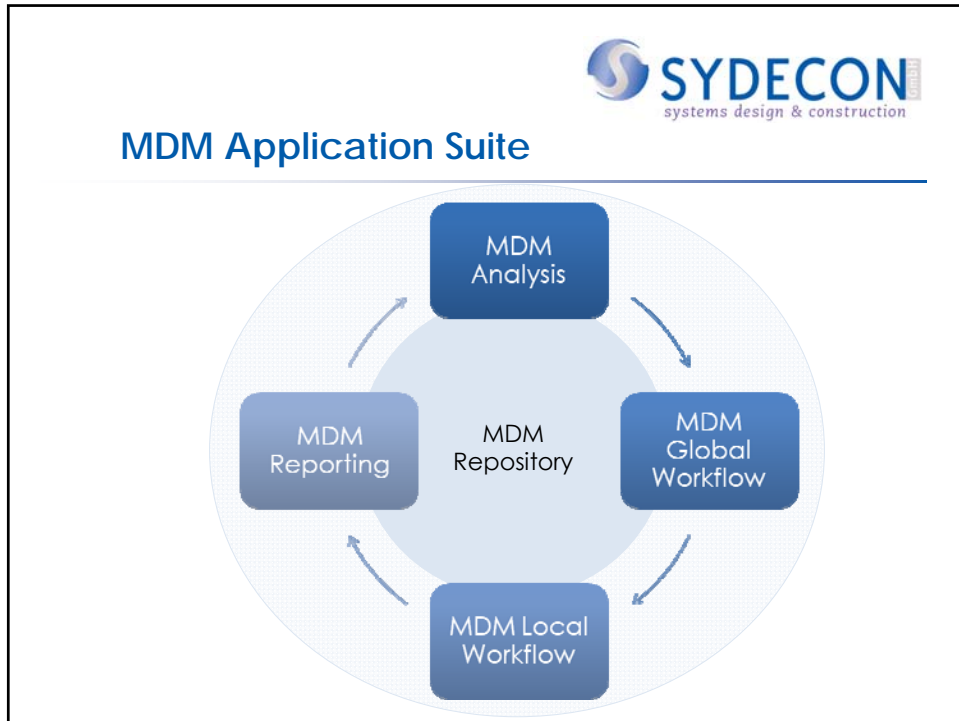


Conclusion

Increased MD quality ...

- ... reduces costs
- ... reduces risks
- ... increases transparency (reporting / analysis)
- ... speeds up the supply chain
- ... observes the laws and restraints

Better MD quality makes your CEO sleep better!





MDM Analyzer Technology

MDM Analyzer ...

- ... is an analytical system for master data
- ... allows rules-based analysis
- ... supports local and global rules
- ... supports local and global responsibilities
- ... allows views restricted by access rights
- ... bases on data warehouse technology
- ... is a centrally installed and maintained system

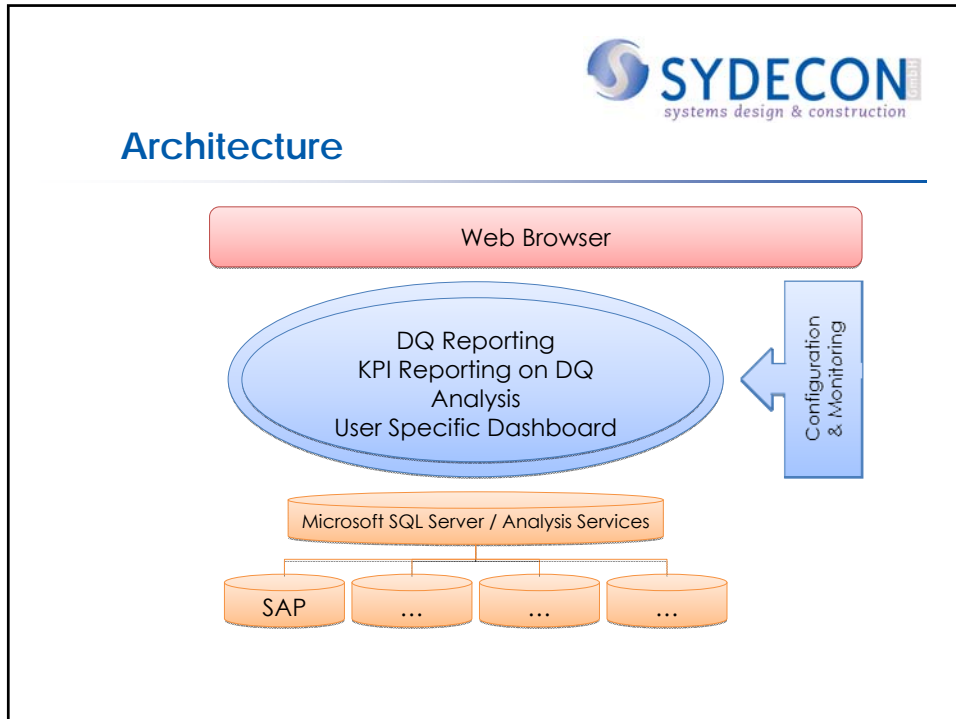


MDM Analyzer Business

MDM Analyzer ...

- ... shows MD quality indicators on high and detailed level
- ... shows the increase and decrease of MD quality over time
- ... finds invalid master data and helps to initiate correction processes in MDM Enterprise Workflow

MDM Analyzer supports MD governance
and controlling tasks in the complete MD organization



SYDECON
systems design & construction

Basic Principles of MDM Analysis

- Administration Areas
An administration area allows the delegation of responsibilities to area administrators and the mapping between the MDM organization and the MDM Analysis system.
- Sources and Rules
Every rule is based on a source. A source defines the data set that is analyzed by a rule:
$$\text{Rule Data Set} \subseteq \text{Source Data Set} \subseteq \text{Database}$$



Rules based Analysis

- There are other ways to detect incorrect data in IT systems, e.g. statistical methods, neuronal networks, etc.
- The rules-based approach is used because the requirements on MD quality are based on rules.
 - MD is used in systems where transactions require well specified information
 - Legal and financial restraints are well specified and define rules
 - Customer and vendor requirements are rules based



Rules and Localization

- Rules can be global
 - Required by a global IT system
 - Based on global company policies
 - ...
- Rules can be local
 - Local laws and restraints
 - Local customer or vendors requirements (e.g. by local logistics providers)
 - ...



What can be analyzed?

Sample rules:

- Customer names should start with a capital letter
- European customers need to have a VAT number
- For each vendor at least one contact needs to be assigned
- Ordered products should have status 'in process'
- For each material in SAP marked as dangerous good, security instructions must be available in the fire department system
- ...



Source Management

Manage sources

Source: 200 Supplier table with Purchase info

Name: 200 Supplier table with Purchase info

Description: Supplier table with Purchase info (LFA1, EINA, EINE)

Source table name(s): LFA1, EINA, EINE

Data specification: LFA1.LIFNR = EINA.LIFNR AND EINA.INFNR = EINE.INFNR

Source type: Global scope and iterable

Distinct object: Distinct object Table name ID column name

KDS assignment:

KDS type	KDS table name	KDS ID column name	KDS value
Country	LFA1	LAND1	
=select item			

Original database tables

Joins

Distinct counting

Access restrictions

Show source history

Clear form Validate SQL syntax Save Save as new Delete

SYDECON
systems design & construction

Rules Management

Manage rules

Process: 06 Customer Management [Edit Process Assignment](#)

Rule: 300 Customer ICC with wrong Rec.account [Search rules](#)

[Show source meta data](#)

Name: 300 Customer ICC with wrong Rec.account

Description: 300 Customer ICC with wrong Rec.account

Detail display columns:

SQL Name	Display Name	Sort Order	
KNAL_KUNNR	KNAL_KUNNR	1	Edit Delete
KNBL_AKONT	KNBL_AKONT	5	Edit Delete

Source: 300 Customer Master Comp Code and Sales Data

Local KDS identifier:

KDS type	KDS value
Sales Organization	RUS0 Browse

[Show source definition](#)

Rule specification: KNAL_KTOID='ICC' and KNBL_BUKR='0182' and KNBL_AKONT > '0000300000'

Area: Area East Europe (Russia/Ukraine)

Result type: Critical Error

KPI rule: Add KPI Calculation

[Assign workflows](#)


[Show value lists](#)

[Show rule history](#)

SYDECON
systems design & construction

MDM Analysis Results

- Dimension Panel in order to navigate in MDM dimensions:
 - Processes, Sources, Rules
 - Areas
 - Key data structures
- Different types of result tables and graphs
 - Current day
 - Compare with history
 - Drill down
 - Drill trough to source data
- Integration
 - Download of error records
 - Integration of MDM Enterprise Workflow




SYDECON
systems design & construction

MD Quality Analysis

List of KPI with target and trend

Welcome Kai-Uwe Baryga



Portal Reports Admin

Filter bar: _____

Settings: _____

KPI numbers of current day

Date: 2008 - Mar - 18

KDS Name	Number Found	Number Total	Percentage of Total	KPI-Value	KPI-Target	KPI-Target Met	KPI-Trend
100 Number of material 01 - Spare Parts	0	710,974	0.00%	5.00	3.00		→
100 Validation Class & Account Assignment	5,769	710,974	0.81%	4.47	3.30		→
300 Customer Name First Letter Capital	355	42,729	0.83%	4.79	3.80		→
300 Customer Name First Letter Capital Counter	42,729	42,729	100.00%	5.00	4.00		→
300 VAT registration number - non blank	2,927	42,729	6.85%	4.20	4.00		→
District Name First Letter	76	42,729	0.18%	4.79	3.00		→
Language for Brazil	774	2,312	33.48%	1.00	2.00		→

Evaluate Add to favorite PDF Export

Country

Company Code

Purchase Organization

Sales Organization

Plant

Area


Process

Source

Rule Result Type

Dimension panel

Show or download error records




SYDECON
systems design & construction

MD Quality Analysis

Compare with last week

Welcome Kai-Uwe Baryga



Portal Reports Admin

Filter bar: _____


Settings: _____

Comparison previous week

Date: 2008 - Mar - 18

KDS Name	Current Day			Previous Week			Comparison	
	Number Found	Number Total	Percentage of Total	Number Found	Number Total	Percentage of Total	Number	Percentage
100 Number of material 01 - Spare Parts	0	710,974	0.00%				710,974	0.00%
100 Validation Class & Account Assignment	5,769	710,974	0.81%				710,974	0.81%
300 Customer-duplicates Test TP	3,529	42,729	8.24%				42,729	8.24%
300 Customer Name First Letter Capital	355	42,729	0.83%				42,729	0.83%
300 Customer Name First Letter Capital Counter	42,729	42,729	100.00%				42,729	100.00%
300 VAT registration number - non blank	2,927	42,729	6.85%				42,729	6.85%
300 test	8	52,888	0.01%				52,888	0.01%
300 test 3	111	710,974	0.02%				710,974	0.02%
Check Customer duplicates	105,064	147,793	71.09%				147,793	71.09%
City Name for Brazil ?	5	2,312	0.22%				2,312	0.22%
Customer ZIP code Reg Ex	33,301	42,729	77.94%				42,729	77.94%
District Name First Letter	76	42,729	0.18%				42,729	0.18%
Language	3,302	42,729	7.71%				42,729	7.71%
Language for Brazil	774	2,312	33.48%				2,312	33.48%
Language for Sweden	2,608	2,608	100.00%				2,608	100.00%
Language Text from global source	41,163	42,729	96.34%				42,729	96.34%
Material Weight	32,565	100,000	32.57%				100,000	32.57%
House Test Rule Street Length BE	1,459	42,729	3.41%				42,729	3.41%
Region	0	42,729	0.00%				42,729	0.00%
RLA for Capital EQ and Brazil - Normal	888	51,617	1.72%				51,617	1.72%
RLA for Capital EQ and Brazil - Urgent	1,012	51,617	1.96%				51,617	1.96%
RLA for Capital EQ and Italy - Normal	211	51,617	0.41%				51,617	0.41%
RLA for Capital EQ and Italy - Urgent	76	51,617	0.15%				51,617	0.15%
RLA for Pack Mat and Brazil - Normal	1,804	51,617	3.49%				51,617	3.49%
RLA for Pack Mat and Brazil - Urgent	1,347	51,617	2.61%				51,617	2.61%
RLA for Pack Mat and Italy - Normal	537	51,617	1.04%				51,617	1.04%
RLA for Pack Mat and Italy - Urgent	537	51,617	1.04%				51,617	1.04%
RLA for Tech Sales and Brazil - Normal	1,499	51,617	2.90%				51,617	2.90%
RLA for Tech Sales and Brazil - Urgent	2,302	51,617	4.46%				51,617	4.46%
RLA for Tech Sales and Italy - Normal	406	51,617	0.79%				51,617	0.79%
RLA for Tech Sales and Italy - Urgent	61	51,617	0.12%				51,617	0.12%
Street Length	1,459	42,729	3.41%				42,729	3.41%
test	0	42,729	0.00%				42,729	0.00%
test 2 KDS	291	100,000	0.29%				100,000	0.29%
Text group by	1,841	76,644	2.40%				76,644	2.40%
Text local	0	2,312	0.00%				2,312	0.00%
Text timestamp rule	0	99,120	0.00%				99,120	0.00%
Vendor classification existence	99,120	99,120	100.00%				99,120	100.00%
ZIP Code for Brazil	630	2,312	27.25%				2,312	27.25%

Evaluate Add to favorite PDF Export



MD Quality Analysis Drill Down

Filter bar: Country: Base: Add to list


Comparison report

	19102007			19102007			Comparison		
Country	Count	Total	Percent	Count	Total	Percent	Count	Total	Percent
AE	0	511	0.00%	0	511	0.00%	0	0	0.00%
AF	0	21	0.00%	0	21	0.00%	0	0	0.00%
AG	1	161	0.62%	1	161	0.62%	0	0	0.00%
AH	0	83	0.00%	0	83	0.00%	0	0	0.00%
AI	0	31	0.00%	0	31	0.00%	0	0	0.00%
AO	0	133	0.00%	0	133	0.00%	0	0	0.00%
AR	154	5,222	3.52%	154	5,222	3.52%	0	0	0.00%

Customer Autocomplete

Customer Name First Letter	Count	Total	Percent	Count	Total	Percent	Count	Total	Percent
C	0	740	0.00%	0	740	0.00%	0	0	0.00%
D	82	740	12.32%	82	740	12.32%	0	0	0.00%
E	82	2,884	3.09%	82	2,884	3.09%	0	0	0.00%

27. Mai 2008



MD Quality Analysis Initiate MD Change Request

Welcome Cornelius Wolf

Portal Reports Admin

Error records: Customer name first letter as capital

Column display mode:

Customer Number	Address Number	Name 1	Country	Language	MDM Requests
0000010260	112011404	Free to be used	NO	E	<input type="checkbox"/>
0000017111	1120047889	Reinhold Quastel	SE	E	<input type="checkbox"/>
0000090117	1120024408	Steen-Hardt, Zweig-Gründe-Fabrik	DE	E	<input type="checkbox"/>
0000090289	1120124214	apexman adw	DE	E	<input type="checkbox"/>
0000090336	1120166494	hauckel stromer gmbh	DE	E	<input type="checkbox"/>
0000090402	1120161165	gsm-200	DE	D	<input type="checkbox"/>
0000090466	1120189998	WOLFF Leasing GmbH	DE	D	<input type="checkbox"/>
0000090771	1120024611	WOLFF Leasing GmbH	DE	D	<input type="checkbox"/>
0000090830	1120166243	WOLFF Leasing GmbH	DE	D	<input type="checkbox"/>
0000094652	7000008353	beckert dealer NEU	DE	D	<input type="checkbox"/>
0000094891	7000008840	Wolfsberg	DE	D	<input type="checkbox"/>
0000094981	7000009018	Wolfsberg	DE	D	<input type="checkbox"/>
0000095010	7000009101	Wolfsberg	DE	D	<input type="checkbox"/>
0000099025	7000009283	Wolfsberg	DE	D	<input type="checkbox"/>
0000110943	1120285237	Wolfsberg	PL	E	<input type="checkbox"/>
0000141135	1120281536	Wolfsberg	NL	E	<input type="checkbox"/>
0000149338	1120287803	Wolfsberg	NL	E	<input type="checkbox"/>
0000170992	7000007239	Wolfsberg	GB	E	<input type="checkbox"/>
0000171137	7000007287	Wolfsberg	GB	E	<input type="checkbox"/>
0000176018	7000007302	Wolfsberg	GB	E	<input type="checkbox"/>
0000176057	7000007365	Wolfsberg	GB	E	<input type="checkbox"/>
0000176059	7000007381	Wolfsberg	GB	E	<input type="checkbox"/>
0000176070	7000007376	Wolfsberg	GB	E	<input type="checkbox"/>
0000176076	7000007394	Wolfsberg	GB	E	<input type="checkbox"/>
0000176078	7000007387	Wolfsberg	GB	E	<input type="checkbox"/>
0000176080	7000007396	Wolfsberg	GB	E	<input type="checkbox"/>
0000176107	7000007421	Wolfsberg	GB	E	<input type="checkbox"/>
0000176116	7000007442	Wolfsberg	GB	E	<input type="checkbox"/>
0000176120	7000007435	Wolfsberg	GB	E	<input type="checkbox"/>
0000176131	7000007467	Wolfsberg	GB	E	<input type="checkbox"/>
0000176134	7000007461	Wolfsberg	GB	E	<input type="checkbox"/>

Link to MDM Enterprise Workflow Process



SYDECON
systems design & construction

MDM Enterprise Workflow Initiated by MDM Analyzer

Save Request Validate Form Submit Request

Change Vendor Master Data

General Request Data

Standard Data

Requester Name: Kai-Uwe Baryga

Title for Inbox: Change vendor due to missing VAT number

Request Urgency: Normal

Comments:

General Data

0. Initial Screen

Vendor: 0005020039

Company Code: 0120 TP International S.A.

Purchase Organization: 0L32 TP Eq,Local Pur.Org

Account Group: LHM Locally Managed Vendors

1. Address

Name: Sydecon Systems Design & Constructi

Name 2: GmbH

Name 3:

Name 4:

Street: Hans-Ullrich-Ring

Street 4:

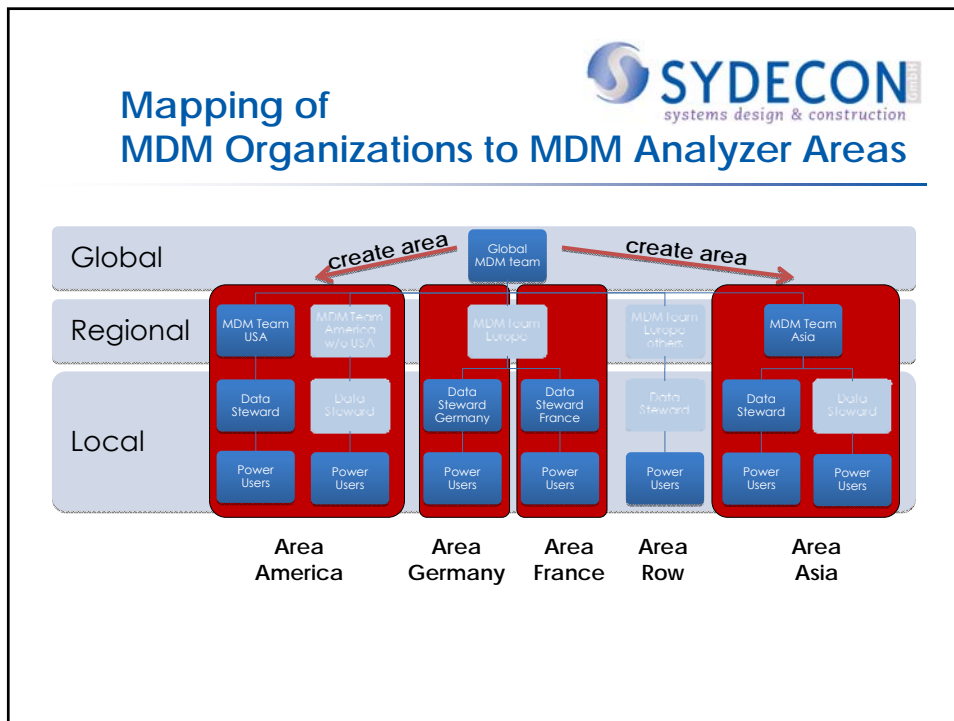
Street 5:

Postal code: 82515

City: Wolfratshausen

Country: DE GERMANY

Data preloaded from ERP System (SAP)



SYDECON
systems design & construction

Area Management

Areas: Area East Europe (Russia&Ukraine)

Name: Area East Europe (Russia&Ukraine)

Description: Area East Europe (Russia&Ukraine) administrated by Kai-Uwe Baryga

Area administrator assignment

Available users

Assigned area admins
Baryga, KaiUwe

Clear form Save Delete ?

Area

Area Administrators


SYDECON
systems design & construction

Use Best Practice

Area administrators ...

- ... often have excellent business experience
- ... have to 'feel' the results of insufficient master data quality
- ... do their best to get perfect master data quality in order to reduce effort for corrections

This leads to best practice solutions that can be reused and delivered centrally



Best Practice

Setting in rules

Manage rules

Process: 06 Customer Management [Edit Process Assignment](#)

Rule: 300 Customer ICC with wrong Rec. account [Search rules](#)

Name: 300 Customer ICC with wrong Rec. account [Show source meta data](#)

Description: 300 Customer ICC with wrong Rec. account

Detail display columns:

SQL Name	Display Name	Sort Order		
KNAL_KUNNR	KNAL_KUNNR	1	Edit	Delete
KNBLAKONT	KNBLAKONT	5	Edit	Delete

Source:

Local KDS identifier:

Area: Brazil

Result type: Error

KPI rule: [Edit KPI Calculation](#) [Remove KPI Calculation](#)

[Assign Workflows](#)

[Show value lists](#)

[Show rule history](#)

Area: [Area](#)

Result type: Critical Error

KPI rule: [Add KPI Calculation](#)

[Assign Workflows](#)

[Show value lists](#)

[Show rule history](#)

Global Administrator can make rules "global"



Access Rights

- Access rights base on the MDM areas and KDS
- Key Data Structures (KDS) are business object related organizational structures, e.g. sales or purchasing organizations, countries and plants
- A user can see data of global rules and rules of his area, if he has the KDS assigned that is required for the data
- E.g. in order to see Brazilian customers with an incorrect address, the user needs to belong to the area South America and have access to the sales organization BR00

Access Rights Setting in Rules

SYDECON
systems design & construction

Manage rules

Process: 06 Customer Management
Rule: 300 Customer ICC with wrong Rec.account

Name: 300 Customer ICC with wrong Rec.account
Description: 300 Customer ICC with wrong Rec.account

SQL Name	Display Name	Sort Order	
KNAL.KUNNR	KNAL.KUNNR	1	Edit Delete
KNBL.AKONT	KNBL.AKONT	5	Edit Delete

Source: 300 Customer Master Comp Code and Sales Data
Local KDS identifier: KDS type: Sales Organization, KDS value: RU60

Rule specification: 'KNAL.KTOKD='JCC' and KNBL.BUKRS='0182' and KNBL.AKONT > '0000300000'

Area: Area East Europe (Russia/Ukraine)
Result type: Critical Error

Callouts:
- KDS assignment: Points to the 'KDS type' and 'KDS value' fields.
- Area assignment: Points to the 'Area' dropdown menu.

Access Rights Setting for Users

SYDECON
systems design & construction

Manage users

Users: Baryga, KaiUwe

Last name: Baryga, First name: KaiUwe
Login name: TP1SEBARYGAK, E-Mail: Kai-Uwe.Baryga@sydecon.de

Column display mode: Surname

User Roles

Available user roles	Assigned user roles
KDSQualityAdministrator	MEMAreaAdmin MEMReportingAdmin MEMReportingUser

KDS Assignment

Available countries	Assigned countries
	AD, AE, AF, AG, AI

Available company codes	Assigned company codes
	0001 SAP Internal 0102 TP Packaging Sol. S.p.A. 0103 Tetra Pak Europe S.A. 0107 Tetra Pak (Canada) S.A. 0110 TP Technical Service Asia 0110 TP Technical Service M.E.

Available purchase orgs	Assigned purchase orgs
	0001 Einkaufsorg. 0001 AES1 RDC GME Pur. Org IAR31 TP Argentina Pur Org

Callouts:
- Application role: Points to the 'Assigned user roles' list.
- KDS assignments: Points to the 'Assigned countries' list.
- Area assignment: Points to the 'Assigned company codes' list.



KPI Reporting

General


- Users can specify KPIs and KPI groups based on existing rules
- Multiple KPIs per rule can be defined
- KPIs are shown in main analysis page
- KPIs can be shown in the KPI dashboard of MDM Analyzer and in the 'Welcome Page' of the MDM portal.
- A user can specify which KPI he wants to use in his dashboard on an individual base



KPI Reporting

Rules Result Normalization


- The results of rules can be normalized based on the absolute number of errors found or the error percentage
- Normalized Rule result = $f(\text{rule result}) \in [0, 1]$
- The function f can be
 - linear, based on the percentage
 - A step function based on values or percentages




KPI Reporting

Result Weighting

- The normalized values can be weighted to a grade between 1 and 5
- In addition a target can be assigned to a rule. A target is also a value between 1 and 5.

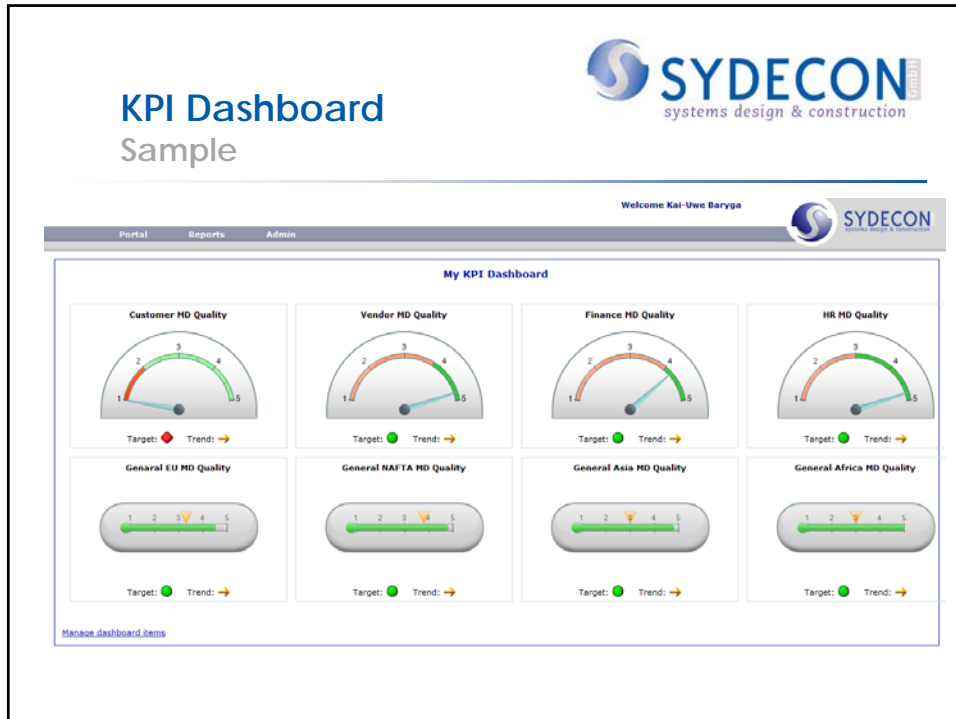




KPI Reporting

Grouped KPI

- KPIs can be grouped
 - A KPI group consist of one or more KPIs
E.g. Customer KPI group can consist of a Customer Address KPI, a Customer Finance KPI, etc.
 - Each KPI in a KPI group can be weighted with percentages.
A more important KPI will have a higher percentage than a less important
 - The user can drill down from KPI groups to the individual KPIs in his dashboard



Performance Aspects

Critical performance areas

- Data loads from ERP systems to the staging area database (less critical than in transactional data warehouses but still a point)
- The Source should not join too many very large tables (we currently use tables with > 250 million records)
- Multiple KDS assignments should be handled with care (no problem in case of real hierarchies, unfortunately this is not always found in SAP environments)
- Drill trough with many records

After processing
the analysis does not cause high loads!



Technology

- Microsoft SQL Server 2005
 - Database Service
 - Analysis Services
 - SQL Server Integration Services (SAP connector)
- Microsoft Internet Information Server
 - Microsoft .NET
 - Developed in C#
- SYDECON is a 



Summary

- The support of central and local analysis is essential for MD Quality control
- Rule based analysis allows direct implementation of business and governance requirements
- Central and local management of rules and access rights make the system more efficient and effective
- Central systems management allows global control of system resources
- Integration in MD maintenance systems adds additional benefits



Ongoing Development

- Closer integration of repository and analysis
 - Specify rules with data specification in repository
 - Specify relationship in repository that can be used in analysis
- Additional automated tasks based on analysis results
 - Mass maintenance
 - Automated data correction



Discussion & Demo