Sponsored by ORACLE

Hub Solution Designs

All the Ingredients for Success: Data Governance, Data Quality and Master Data Management

Hub Designs, 188 Whiting Street, Suite 6A, Hingham, MA 02043
(781) 749-8910, www.hubdesigns.com, hubdesignsmagazine.com

Table of Contents

1. Introduction

Improving data quality is one of those timeless things which can provide value on its own, or it can be done as a first step towards something else such as master data management (MDM), or it can be done together with MDM.

This white paper will explore all three of these use cases:

1. **data quality independently**, which is good for cleaning data inside an application or as it's being moved into another application
2. **data quality as a first step towards MDM**, which allows you to start with one application knowing that MDM will be introduced as more applications get into the act
3. **data quality in parallel with an MDM implementation**, which is good for cleaning up and standardizing data across multiple applications

We'll try to get you to see data quality not just as an "add on", but as an equal component of an enterprise information management or master data management initiative. And we'll examine the role that data quality tools plays vs. the role an MDM hub plays.

I always suggest starting with the business pain or problem, not "the solution". In a lot of situations, it turns out that the lack of the "Single Customer View" is really the lack of the "Reliable Customer View". When that's the case, a strong data quality solution may be all that is needed, or companies in that situation may need to go on to the "data quality plus MDM hub" use cases mentioned above.

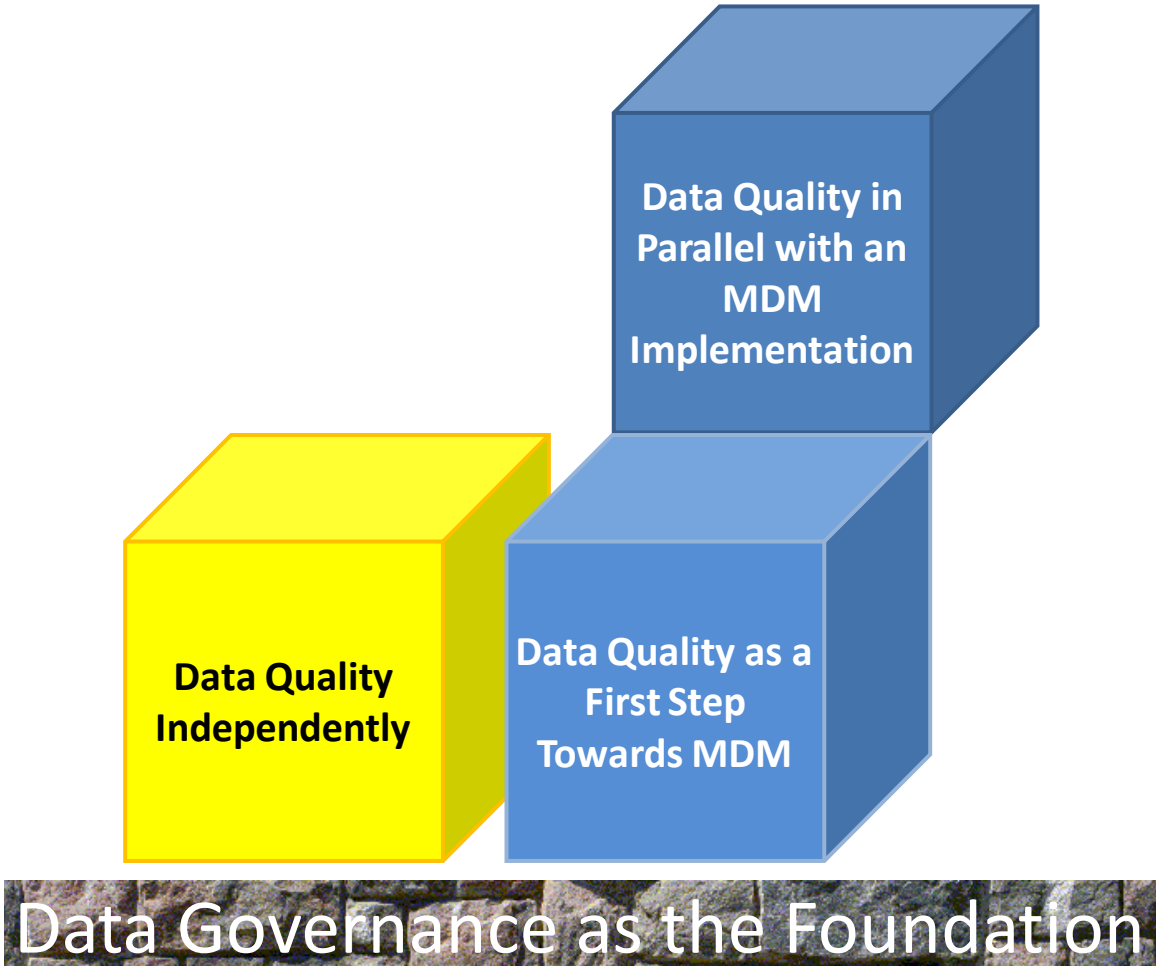
We'll review the elements of a robust data quality and master data management strategy. The most important thing is to make sure you're setting out to solve real and significant business problems.

Many times, there are "presenting symptoms" of business pain that people are complaining about, such as:

- **bad decisions** made in the past, based on reports with wrong or missing information,
- the **costs of duplicate data** in your customer master database,
- cash flow being negatively impacted because **invoices are going to the wrong customer addresses**,
- supplier master issues causing millions in **avoidable costs due to lost volume discounts**,
- penalties incurred for **not complying with regulations** on customer data and privacy,
- information-heavy projects that **run over time or require lengthy rework** after go-live,
- aging systems that are **nearly impossible to decommission** because clear retention and destruction policies have not been defined,
- IT staff spending a lot of time **re-integrating data and dealing with "data fire-drills"** because data service level agreements and "fit-for-use" levels have not been established and tracked.

All of this traces back to lack of data governance and poor quality data in the end. Master data management technology can address a lot of these issues, but only when driven by an MDM strategy that includes a vision that supports the overall business and incorporates a metrics-based business case. Data governance and organizational issues must be put front and center, and new processes designed to manage data through the entire information management life cycle. Only then can you successfully implement the new technology you'll introduce in a data quality or master data management initiative.

2. Which Approach is Better?



I sometimes compare the three data quality and MDM use cases depicted above to a person being left handed or right handed.

The "left handed" approach is data quality independently. Data quality as a first step towards MDM is the "right handed approach". Data quality in parallel with an MDM implementation is folding left and right hands together.

Which is better? The answer obviously depends on your requirements. But it's good to have choices.

Choice gives you the flexibility to arrive at the right data quality and MDM strategy for your company – not just the strategy that happens to fit best with the products of the MDM vendor with whom you happen to be talking.

Some MDM products, because they have a built-in data quality tool, limit you to either using their built-in tool, even though it's far from being best-in-class, or essentially deactivating it and buying a best-of-breed data quality tool and grafting it onto your hub, an integration exercise you'd probably rather not have to do.

But having data quality and master data management separate but tightly integrated allows you to "mix and match" gracefully.

2.1 Data Quality Independently

In some cases, simply adding a data quality tool is all that is required.

My client base has included many clients where they have struggled with an enterprise resource planning (ERP) or customer relationship management (CRM) application that didn't have any built-in data quality capabilities. In cases like that, even though my firm specializes in master data management solutions, we recommended a more straightforward approach, to add a top notch data quality tool to their IT arsenal, since it would address many of the presenting symptoms the client was telling us about.

Another scenario that comes up frequently is a business intelligence or data warehousing application, where the data being extracted from the operational systems is – after some basic analysis – too "dirty" to bring into the warehouse in a useful way. We've all seen those types of "data horror stories" where lax application data entry standards have allowed less than ideal data into the transactional application.

One of the most basic rules of data management is that data is usually fine in the first system or context where you enter it or use it. But when you try to extract it from that system and use it more broadly, in this case an ERP system, all the warts are revealed. That's where a data quality tool comes into play.

Don't fall prey to the old fallacy that "we can fix this with a few lines of SQL code". That's a slippery slope. Once you start down that path, you're committing to creating your own "in house" data quality tool, and as you convert more business units into the ERP system, you'll have to take their peculiarities and warts into account too.

Better to recognize from the beginning that you've got a data quality problem here, and that there are high caliber tools designed specifically to solve that problem. Not only that, those tools have been around for a long time, have been highly refined by now, and are available from major vendors who do a great job supporting them and integrating them into their other products. That sounds like a clear case for a "buy" decision vs. a "build" decision to me.

Once you've gotten past that decision, implementing data quality tools independently has become much simpler in the last few years. Nowadays, you should be thinking enterprise-wide from the beginning. Set up your data quality tool as an enterprise-wide web service that can be invoked from any application. That way, as you adopt it in your first application like the ERP system example we're using here, you get your planned return on investment (ROI).

But then, when you get some time in your IT resource calendar, you can implement it more broadly as well. Your web site might be a good candidate for the next adoption point. If you allow customers to create accounts on your web site, make sure that the account details such as the address they're giving you are a valid, deliverable address.

As we move towards a more self service, web driven environment, data quality tools need to be integrated into every step of these processes. Otherwise, all we've done is given our customers, our vendors and our employees the ability to accidentally (through typos) or deliberately (through fraud) gum up the works of our business processes and, in the end, hurt our bottom lines.

2.2 Data Quality as a First Step towards MDM

This also comes up quite a bit. Some companies aren't ready to plunge directly into a large MDM implementation, or they will take on a master data management initiative, but over a fairly long time frame.

In this case, I usually look to see what value can be unlocked in the short term (in three months or less) via a data quality tool.

In a lot of cases, there's an "800 pound gorilla" in the business – an area of business pain that's so bad, people have stopped talking about it. People have tried multiple times to tackle it and failed (usually with the wrong tools). If it's a problem that a data quality tool can help solve, even when an MDM initiative is on the horizon, I recommend that, because it's an excellent first step, and will help develop some of the discipline needed for the later MDM initiative.

The company will still need to develop its vision for the data quality initiative, and create a DQ strategy that supports that vision. A business case will still be needed, if for no other reason so that you'll know at the end of the project whether you achieved your expected ROI or not.

A beginning towards the data governance framework, organization and processes will also be needed. None of these steps can be skipped just because you're doing a data quality project rather than a full blown MDM project.

Don't forget to have the business driving the initiative. Data quality initiatives driven by IT typically fail. Only the business people in the company can make the decisions about when data is "good enough" (i.e. fit for purpose). Although IT sometimes resists and wants to lead these initiatives, and even though the business team sometimes regards anything having to do with computers, databases, and information as being "IT's responsibility, right?", don't make the mistake of doing a data quality or MDM initiative without the business being firmly in the driver's seat.

That balance of "business and IT alignment" – where the business takes the lead but has strong support from and aligned leadership from IT – is a very difficult balance to strike, and may require a new model of engagement at your company. I recently had a very revealing conversation with a Global 2000 CIO, who grasped the implication of this shift, and who said "I'm going to have to start hiring more artists and fewer programmers."

He saw that as his company moved towards a business driven, enterprise-wide data governance model with a data quality initiative leading towards multiple MDM hub implementations, that IT's role relative to the business was going to change, and he embraced that change.

Data Quality as a first step towards MDM can be a powerful strategy, because it allows you to "ease into" MDM in manageable phases, and get distinct business value at each stage. "Think big, start small" has become a piece of conventional wisdom in the MDM world, and this is a great way of doing that.

A good data quality tool is a great building block, and you may already have implemented one in your organization that you can now adopt more widely, as mentioned in the previous section. A data quality tool previously implemented for use in the customer data warehouse can become a core building block of your enterprise-wide data governance efforts, even before your MDM hubs are implemented.

2.3 Data Quality in Parallel with an MDM Implementation

This is probably the most common use case that I see, since my firm specializes in MDM, data governance and data quality. It's important not to think of this as a "big bang", since the current generation of MDM hubs have all been designed to integrate tightly with a data quality solution.

Most hubs have a built-in data quality tool, and unfortunately a lot of them are not very good. One of the criteria I use when evaluating an MDM hub with a client is how good its built-in data quality capabilities are, and how important those capabilities are, relatively speaking, to the client. If the client's requirements are weighted very heavily towards needing a strong data quality solution, which will shape the evaluation of the MDM hub accordingly, sometimes eliminating an otherwise likable hub with a data quality component that just isn't robust enough.

This use case is sometimes the only one people consider, but hopefully the previous sections have got you thinking about the "Data Quality Independently" and "Data Quality as a First Step towards MDM" use cases.

This is still a good time to take the approach of setting up the data quality tool as an enterprise-wide data quality service, with the MDM hub as its first "customer". That way, the business rules and services you create for the hub will immediately be available for use outside of the MDM initiative if that's appropriate, and again, you'll be able to claim even wider adoption and larger than anticipated ROI, if you are successful in finding other points of adoption elsewhere within the enterprise.

And the data governance approach is even more critical here, because now your newly constituted data governance organization will be making decisions not only about data quality for the enterprise, but also about how to implement one or more master data management hubs. Talk about a trial by fire! But it is good practice to have a data governance organization participate in implementing the MDM hub, because after the hub goes live, the data governance team – from the executive level, to the Data Governance Office, right down to the individual stewards – are the people who are going to be responsible for using the hub, maintaining it, keeping it funded, promoting it to the rest of the enterprise, plumbing it up to all of the other source systems in the enterprise's architecture, and gradually changing the processes and culture of the company to be centered around one or more master data management hubs as the "Single View" of the critical information assets of the company.

Trust me, that's not a short term process, but it all starts with – in this case – a well-chosen data quality tool and MDM hub. Just don't forget the commitment you're making to a balanced MDM program, with a shared vision and strategy, a data governance organization, the appropriate business case and processes, in addition to the requisite data quality and master data management technology.

And to be successful, recognize the importance of organizational change management – these projects require significant changes in how business and IT relate, and in how your company thinks about, manages and invests in information.

3. The Importance of Data Governance

3.1 Business Ownership is Essential

There's a natural tendency on the part of business people to assume that because master data resides in computers and databases and sounds technical, that it must be IT's responsibility to take care of it.

But nothing could be further from the truth. The reality is that while IT people are certainly involved, they'll never be close enough to the data to be able to tell a prospect from a customer, or a good address from a bad address.

Not only that, but if IT allows business to shift the ownership of data governance to IT, then IT will own it forever. Some people in IT might not think that's a bad thing, but one of the problems with that is that the business will then see data governance as "Someone Else's Problem" (SEP). This can make it very hard to get the recurring funding needed for data governance.

On the other hand, when business is fully engaged, is out front leading the data governance initiative and feels a real sense of ownership, IT can still be heavily involved in a supporting role, but with the business leading the charge, funding is usually not a problem, resources are made available when needed, and most of the issues that typically plague data governance programs are less of a problem.

This model, with business leading and IT supporting, is much healthier, but it requires a good working relationship between business and IT, and may require some remedial "business / IT alignment" work to rebuild the partnership between the business and IT. The reality is that it's going to take a tightly integrated team of business and IT people to build and manage the MDM "stack" and the data governance organization that will be using it.

3.2 The Business Decides What Data Quality Means

The business data stewards will handle the data governance, but they'll need a lot of support from IT. There will be various user interfaces and workbenches to carry out the business decisions the data stewards make.

They'll interact with workflow queues, data quality front ends, business intelligence scorecards, rules engines, and so on. The average data steward will make the "power user" of the past look like an amateur. But in the end, it's the knowledge of the business that makes the data stewards valuable.

And that's why most companies fill these roles internally. They want people from the Order Management department, or from Customer Service, Sales Operations, or Finance. The company needs people with intimate knowledge of the industry and the organization and its customers, products, suppliers, employees, locations, geographies, etc.

Many of these business data stewards will not be full time dedicated positions. They will be people who are doing this job already as part of their existing responsibility, but that arrangement is now being made more formal, and their past hard work is being recognized. Hopefully, their attention to detail and "getting the data right" will now become part of their annual compensation and bonus plan, in recognition for their taking on formal data stewardship responsibilities.

The **Data Stewardship Teams** in various functional areas and geographies around the business typically report up to a Data Governance Office, which typically does have dedicated personnel.

This **Data Governance Office**, which functions similarly to a Program Management Office that coordinates multiple ongoing programs and projects in an IT context, provides leadership and coordination to the different Data Stewardship Teams. It also creates data policies, and includes global process owners from the business, as well as a global solution owner from IT. It runs the overall data governance program, coordinates with other groups like Legal and Enterprise Architecture, and provides a single point of contact for all data governance related matters.

Finally, the Data Governance Office usually reports up to an executive group of stakeholders, which are often referred to as a Data Governance Council or **Executive Steering Committee**. This group sets the overall priorities and funding and provides oversight for the entire data governance effort, as well as making some high level policy itself and resolving issues that are escalated to it because they couldn't be resolved at the lower levels. Sometimes this function is taken up by an existing body, and sometimes it's a new group that is pulled together by the executive sponsor who's responsible for forming data governance in the first place.

3.3 IT Handles the Technology (and More)

Although we just described how business-driven the whole data governance effort is, don't despair if you're an IT person – for two reasons. First of all, there's a lot of cool technology needed to support master data management and data governance.

There may be things like:

- The MDM hub itself, providing metadata and hierarchy management as well as audit features
- A data integration tool using a service bus and other service-oriented architecture (SOA)
- A data quality tool providing data profiling, entity resolution and data standardization capabilities
- Security and authentication, perhaps including single sign-on and enterprise identity management
- Analytics, including data warehousing, business intelligence and business activity monitoring

Secondly, most data governance efforts, although they are ultimately business-led (and need to be in order to be successful) are conceived within the IT organization. So there's the pride of authorship – of bringing something new into the company.

It requires a delicate hand – because IT has to learn how to suggest and encourage without controlling. The business has to feel like it is actually running the data governance organization, even though IT had a big part in getting it off the ground.

I call this "leading from the back of the room", and it can be difficult. But it's a form of leadership that IT can learn, like being a power of example to a younger sibling. The business people who are part of the data governance organization will need mentors, and handled correctly, it will lead to a closer business/IT working relationship.

IT has to avoid the temptation to take over the Data Governance Office, or to allow the business to shift responsibility for data governance back to IT. That's why getting agreement up front on a data governance mission statement and charter is so important, and getting the initial design for data governance signed off by the Executive Steering Committee can prevent setbacks when things get tough.

4. Product Data Quality and Master Data Management of Product Data

4.1 Product Data Quality

If you've only worked with customer information in master data management projects, you'll run into some differences pretty quickly when starting to work with product MDM.

Before doing MDM, I did enterprise resource planning (ERP) projects for about ten years, most of them with manufacturing companies with large item masters or product catalogs, so when I started doing product MDM, these differences weren't a surprise for me.

First, product information is, in general, more complex, with more attributes than customer information. And product information is poorly structured. Most of the useful data will be jammed into the "Description" attribute, abbreviated six ways from Sunday, with no logical standards that you will be able to determine.

Secondly, validation is a big issue. With customer data, addresses can be validated against postal standards. In some industries, there has been some progress towards standardizing some elements of product information. But there's still a long way to go.

Thirdly, data volumes can be surprisingly high. Companies with less than 50,000 customers can end up having hundreds of thousands of products. It's called "SKU proliferation" (for Stock Keeping Units), and if you do a quick Google search on that term, you'll find more than 96,000 search results. So it's a big problem for most businesses.

A lot of companies can only deal with creating slight variations of their products by creating a whole new part number, which leads to a lot of essentially duplicate products. Not duplicates you can merge back into one record. But in the new Product Hub, you'll be able to handle variations another way, by creating one product record and recording variations without having to create a whole new part number.

A fourth issue is the number of categories of products, and the fact that the product attributes vary by category. A cell phone's attributes are going to be completely different from a ceramic coffee mug. Even the number of attributes is going to vary. So you're going to need a "taxonomy" to handle the classification of products by category, and your Product Hub is going to need to assign the number, type and default values of attributes based on the position in the taxonomy that a product is assigned.

That's an elaborate way of saying that when you create a new product whose Category = "Cell Phone", the attributes should be automatically specified by the taxonomy, and the default values should be intelligently defined for you, so you (as the product manager) only have to fill in the values that don't equal the defaults.

So this categorization of products into a taxonomy is a major way that product information is different from customer information. Business customers can certainly belong to a corporate family tree, but that doesn't drive attribution the way a product taxonomy does.

The impact on data quality is very direct – if you can determine what product category an incoming product coming in via an interface or feed belongs to, you can align it with that category in the product taxonomy, and then compare the new product's attributes to the standards on file for that category in the taxonomy. Any product whose category can't be determined can be routed to a data steward for manual review.

There are three common methods of dealing with product data quality issues:

- **Manual effort** – this is really only practical if you have a very small number of products or if you're conducting a pilot program, for example. If volumes are larger, as they typically are, this becomes impractical.
- **Custom code** – this is a slippery slope, because initially you think a small SQL or PL/SQL program should be able to handle the problem. But then the variability of the product information catches up with you. There may be dozens or hundreds of different types of products in your portfolio, and they may all have different numbers and types of attributes, and potentially different types of data quality problems. So in the end, you realize you've signed up for something akin to Napoleon's Waterloo (an unwinnable battle).
- **Traditional data quality tools** – most data quality tools are not designed specifically for product information. They're designed for customer data and are not well suited to deal with the unstructured, highly variable nature of product information. So in a way, it's similar to writing custom code. By the time you end up dealing with all of the exceptions and writing scripts in the DQ tools scripting language, you're still going to end up with a maintenance nightmare.

Given the differences between product information and customer information, and the difficulty using manual efforts, custom code or traditional data quality tools to deal with product data quality issues, the best way to tackle serious product data quality problems is with a professional caliber tool.

Look for something designed expressly for handling the complexity, validation requirements, volume and variability of product information. The ability to recognize the context of data is known as "semantic recognition", and it's a key capability to look for, because otherwise, your data quality tool is going to treat every product record in a flat, undifferentiated way.

Instead, look for a tool that can tell from the context of the data what type of product it is. As George Orwell said in "Animal Farm", "All animals are equal, but some animals are more equal than others." That way, the tool can handle the variability of rules for different product categories without manual scripting.

Scalability is a key characteristic too, since as we already discussed, product data volumes tend to be very high. You may need to enforce product standards across thousands of product categories and millions of product records, and do so while providing a product data quality web service handling calls from dozens of internal applications.

Integrated governance capabilities are increasingly important as well, since the business data stewards we mentioned above will be managing product data quality, and will be responsible for overseeing data remediation and exception management. The entire product data quality solution will need to be maintained by business users, with minimal assistance from IT, after the initial implementation is over.

Look for a tool that can easily output data in many different formats – the canonical format needed by your Product Hub, and the many other formats needed by various target systems in the enterprise that you may need to publish to, in a variety of different languages, standards and formats.

The main process flow that the product data quality tool will be part of is: automated data cleansing, followed by matching and duplicate resolution, then either updating the Product Hub if the product record already exists or adding the record to the Product Hub if the record doesn't already exist.

4.2 Product Master Data Management

Of course, it doesn't make sense to try to do Product MDM without an integrated Product Data Quality solution as well. So we'll assume that whatever you're going to do in your Product Information Management strategy, you're going to include both Product Data Quality and Product Master Data Management.

The biggest benefit of doing that is that the Product Data Quality solution can make sure that data flowing into your Product MDM Hub is ready for that hub. Whatever standards were defined by the data governance organization created earlier, Product Data Quality solution enforces those standards on data going into the hub. Think of it as a "data firewall".

And since those standards vary by product category, a robust Product Data Quality solution can save a lot of manual review, since screening incoming data can otherwise be a very labor-intensive process that soaks up a lot of your data stewards' time.

Eliminating "false positives" – data that looks like it violates standards but really doesn't – is a huge savings, because it allows a much bigger portion of the data migration process to be handled automatically instead of being routed to a human being for review.

And making both the Product Data Quality solution and the Product MDM Hub itself available across the whole enterprise via easily invoked web services drives both wider adoption and higher return on investment.

Having a single point of maintenance for Product Quality and Hub standards and governance helps to break down the all-too-common silos, and makes the business process benefits more visible, since the impact when a change in those standards takes place will be readily apparent.

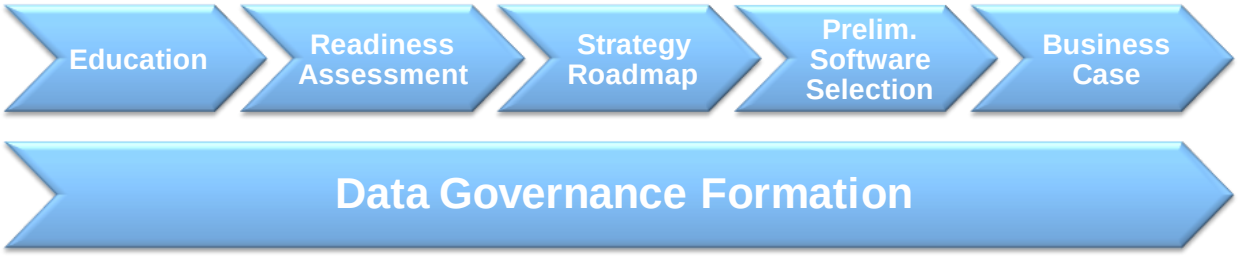
Even having a central place for the management of product attributes is a huge improvement for most companies, since the data fragmentation problem means that most companies don't have a central place to manage product descriptions, pricing, product images, or other elements of their product catalog.

The Product Data Quality and MDM services can be called from any application in the enterprise, so over time their use will become more and more widespread, and the "culture of data quality" will become more widespread as well. Of course, the data governance organization will need to be involved, both in promoting this, and in managing it.

All of this will help drive adoption and increase ROI across the enterprise more quickly, because these additional points of adoption, although they should be documented and attributed to the MDM solution and the data governance organization, are usually not part of the initial business case.

Since they aren't, the effect is to make the initial return on investment even more attractive, and to help increase the Net Present Value of the project even further. In some cases, this can be used to fund later stages of the project.

Putting It All Together – The Strategy



At Hub Designs, we specialize in helping our clients get everyone on the same page through rapid education, then a brief readiness assessment, followed by a Strategic Roadmap and Business Case.

In parallel with that, we help our clients to create an enterprise-wide Data Governance Organization.

Putting It All Together – The Delivery

After selecting your software platform, you'll need a tightly integrated Data Governance Team of business and IT people to implement and manage the Data Quality / MDM "stack".

The business data stewards will handle the data governance, with lots of support from IT. The IT stewards on the Data Governance Team will handle the more technical aspects of the data quality tool and the hub:

- Metadata management
- Security and authentication
- Complex data quality rules, entity resolution and matching
- Discovery of meaning in product data quality
- Data integration (middleware)

The end result of all this, from a Product Information Management perspective, is to be able to:

- Provide higher quality data; prevent mistakes and bad decisions
- Enforce standards across thousands of product categories and many systems & processes
- Increase productivity – IT does not have to write laborious programmatic rules for product data cleansing
- Accelerates time to market and reduce costs for new product development
- Enable automated replenishment programs and digital images for e-commerce
- Facilitate product-related compliance with government regulations (pharma, life sciences, etc.)

This "virtuous cycle" effect, of being able to gradually drive common standards across all systems in the enterprise, of course takes time, but the combination of a Product Data Quality solution and a Product MDM Hub is a powerful one, that will impact both the culture (people), business processes, technology environment and information architecture of the deploying company over time.

Three to five years from now, the internal evangelists who began this transformation cycle will hardly recognize their own company.

5. Customer Data Quality and Master Data Management of Customer Data

5.1 Customer Data Quality

High quality customer data continues to be a key pain point for businesses look to improve operating efficiencies across the enterprise. As we saw in our discussion of product information, customer data is significantly different from product data.

Where product data is largely unstructured and with an attribute set that varies by product type, customer data is usually structured, and the attributes don't vary very much, although having the ability to capture user-defined fields is always helpful.

Customer data can be modeled using a "party" data model, which allows the enterprise to capture individual organizations (businesses) and consumers, as well as their relationships to one another, such as corporate family trees and consumer household relationships. A rich party data model also allows for today's mobile workers, who may work for several companies at once, telecommute from home while also being attached to one or more corporate offices, and have multiple sets of personal and corporate contact information.

This, however, significantly complicates the data quality challenge, because it can present exponentially more data and relationships to be profiled, audited, cleansed, and matched.

Social networks have become a factor in this as well. Most companies now want to keep track of the social network profiles of their customers, which makes it possible to track what those customers are saying about the company's brand on those social networks. But that adds even more data that must be gathered, attached to individual contact records, and tied back to organizations (parties).

So the strategic importance of enterprise data quality capabilities has gone up dramatically in the last few years, not only as a component of master data management, but in its own right as well.

One reason for this is the increasingly interconnected nature of our society. In the 21st century, these events show how the world is simultaneously becoming a more dangerous and a more regulated place:

- the events of September 11th, 2001
- the 2004 Madrid train bombings
- the 2005 London bombings
- the 2008 Mumbai attacks
- the collapse of Lehman Brothers in September 2008
- Bernie Madoff pleading guilty to the largest Ponzi scheme in history
- the bankruptcies of General Motors and Chrysler in June 2009
- the sovereign debt crisis in Europe which started in April 2010
- the recent downgrading of the United States government's credit rating by Standard & Poor's

The terrorist incidents above have led to new regulations like the USA PATRIOT Act of 2001, the Homeland Security and SAFETY Acts of 2002, and the Know Your Customer banking regulations of 2002.

The global financial crisis of 2007-2011 has led to new regulations and legislation like Basel II and Basel III, the American Recovery and Reinvestment Act of 2009 and the Dodd-Frank Act of 2010.

The financial services industry and the intelligence community in particular are continually looking for better ways to "connect the dots".

But organizations in all industries are faced with the necessity to do the following:

- Avoid doing business with sanctioned or risky individuals and entities, and comply with "Know Your Customer" (KYC) and Enhanced Due Diligence (EDD) obligations, and

- Use master data management, data governance and enterprise data quality tools, plus business intelligence and advanced analytics with their enterprise data to deliver business benefits such as:

  - Increase revenue and CRM effectiveness,
  - Reduce costs and improve operating efficiencies and productivity across the enterprise,
  - Streamline compliance, credit and risk management efforts
  - Improve overall agility and decision-making

What are the characteristics of an enterprise data quality tool that can handle product and customer data?

It should be seamlessly integrated with the other components of your environment. "Out of the box" integration is a big win. When you select an MDM hub, and have other components of your architecture like a "front office" CRM platform and a "back office" ERP suite, having an enterprise data quality solution that is pre-integrated with as many of those other products as possible will save you a significant amount of time and money, versus having to do that integration yourself.

It should be a modern, accessible architecture, based on Java and service-oriented architecture (SOA). This will make it easier to create and provision enterprise-wide data quality web services, which are one of the keys to widespread adoption and better than originally expected ROI. Also, having an accessible user interface and multi-user project support makes it easier to support the collaborative, controlled access to the business rules by the data governance team that you'll need.

The product should support delivery of complete solutions and applications, so that you can use it as a platform for developing internal capabilities and web services across the enterprise, including the data quality capabilities needed to support access from your web site by customers and employees, from your CRM and ERP systems, from your MDM hub, and from data migration projects around the enterprise.

A truly end-to-end enterprise data quality solution that can handle both product and customer information is very unusual in the marketplace. Most can only handle customer data, and are fairly specialized to a particular approach, or are based on a fairly antiquated architecture.

Put the products side-by-side, and do a bake-off during the sales cycle with your own data. This is always very revealing when it comes to comparing the results of profiling, standardizing, and matching your product and customer information. Use your actual requirements as your guide. You'll be much better off than if you just compare "features and benefits" between the competing products.

Finally, look for an enterprise data quality solution that fits into an overall product roadmap, with master data management, data integration, and data quality. And if the vendor can provide your front office CRM and back office ERP applications, and your business intelligence and advanced analytics as well, so much the better.

5.2 Customer Master Data Management

Sometimes when I'm talking to a new client who's just getting started on a customer MDM project, I can tell they're wondering "how hard is this project going to be?"

I wish there was a better answer. The reality is that any MDM project (but it seems particularly customer MDM projects) are difficult – in large part because of the internal politics that inevitably break out.

I'm reminded of a couple of quotes by Thomas Edison, who said:

- "Opportunity is missed by most people because it is dressed in overalls and looks like work."
- "Genius is one percent inspiration and ninety-nine percent perspiration."

I wrote an article for *Hub Designs Magazine* in October 2007, in which I rebutted someone who said that because MDM solutions were so complex, they were "too difficult".

In that article, I quoted President John F. Kennedy's famous speech to the country about going to the moon in September 1962:

> "We choose to go to the moon. We choose to go to the moon in this decade and do the other things, not because they are easy, but because they are hard, because that goal will serve to organize and measure the best of our energies and skills, because that challenge is one that we are willing to accept, one we are unwilling to postpone, and one which we intend to win …"

My point is that customer MDM is probably the single most important thing most companies can do to increase their value. If you

- don't have a good handle on your customers' lifetime value,
- don't have accurate customer segmentation,
- don't have visibility to your customers' corporate family trees,
- can't position a Next Logical Product as part of your approach to cross selling,
- can't provide real-time decisioning to target the right person with the right offer at the right time,
- can't incorporate the fast growing amount of data (including information from social networks),
- can't integrate your customer data across the enterprise to provide the Single View of the Customer,

don't worry – you're not alone.

But there is something you can do – start by getting as educated as you can on master data management and data governance.

Then do a Readiness Assessment on your people, processes and technology, using the "4-D Approach" outlined in my article in the May/June 2011 issue of Information Management magazine, or using the Information Maturity QuickScan of MIKE2.0, the open source methodology for Enterprise Information Management.

After the Education and Readiness Assessment phases, you'll be ready to develop your Strategic Roadmap. Hub Designs Magazine published an article by James Parnitzke on how to build a strategic roadmap in March 2011. This has been one of the Top Ten most popular articles on the Hub Designs Magazine site this year.

The steps boil down to:

1. Develop a clear and unambiguous understanding of the Current State (the "As Is")
2. Define the desired Future State (the "To Be")
3. Conduct a Gap Analysis, which leads to closing each Gap with a Gap Closure Strategy
4. Prioritize and arrive at the optimum sequence of Gap Closure Strategies
5. Develop the Road Map, which shows the Gap Closure Strategies over time grouped in initiatives

After developing your Strategic Roadmap, you'll want to create a compelling Business Case. The main components of a business case are typically an attempt to quantify costs and benefits over time. Once you have spelled those out, you can plug them into an ROI model.

The costs and benefits over time will give rise to a cash flow, which will be discounted (using the concept that money today is better than money in the future). That discounted cash flow is known as the Net Present Value (NPV), and NPV is an important measure of the value of the project.

Generally, if your project's benefits exceed your costs by a healthy amount, you'll have a positive ROI, a short payback period (the time required for the project to return the original investment), and a large enough Net Present Value to make the project worth pursuing.

After developing a robust Strategic Roadmap showing why your company should pursue MDM as a business strategy and how it will enable the company to deliver on its overall business strategies, and creating a compelling business case with realistic cost estimates and actual benefit statements based on interviews with business people throughout the company, you'll be ready to select and implement the solution itself.

But how do you select the right MDM solution for your company? Clearly that's not as simple as picking out a new laptop or desktop computer, or the countless other types of technologies that have become commodified over the years.

Here are the steps I recommend for selecting an MDM hub:

1. Don't start by talking to MDM software vendors – **start by learning as much as you can** about master data management and data governance
2. Once you start to better understand the MDM market (bring in a consultant to help you if necessary), start to **create a selection scorecard**, consisting of a list of criteria for MDM vendor and software selection, based on your business requirements.
3. **Assign relative importance (or weights) to each item** on the list of criteria. This allows you to prioritize the "must have" criteria higher than the "nice to have" requirements.
4. After learning as much as you can about MDM and data governance, and gathering your requirements in order to build your list of criteria, and weighting those criteria to reflect your priorities, **then start talking to vendors**. Conduct a Request for Information (RFI) or Request for Proposal (RFP), or simply invite them in for meetings to get to know them and their products, and to allow them to get to know your selection team, your company and your requirements.
5. Finally, **score each MDM vendor's ability to meet your requirements** on your selection scorecard.

You want to take into account the product's functionality and cost of course, but also the vendor's underlying technology architecture and vision, as well as their financial viability and their professional services, training and technical support capabilities.

So what are the underlying capabilities to look for in a customer MDM solution? A master data management solution should, by definition, improve the quality and consistency of master data across the enterprise.

The solution should provide capabilities like the ability to integrate multiple sets of data (for example, customer data from many different source systems) and produce a **single enterprise-wide 360-degree view of the customer** – that is, the **Single View of the Customer** - for example, one “golden record” per customer, which can then be synchronized back out across the enterprise as the official “clean, accurate, de-duplicated and enriched” information on that customer.

The solution should have closed-loop data quality services, so customer and address quality can be systemically improved, providing cleansed, matched, and standardized customer records.

The solution should be tightly integrated with service-oriented middleware, allowing two-way integration between the central hub(s) and source systems in order to acquire, synchronize, and federate data.

The solution should facilitate and support enterprise data governance, so that data can flow seamlessly from source systems to the central hub(s) and back out again, and exceptions can be dealt with by a relatively small number of global data stewards using collaboration tools, workflow and business rules.

The solution should have advanced hierarchy management capabilities, to allow end users to manage customer hierarchies (both corporate family trees and consumer households).

The solution will have a robust security model, so end users can be assigned roles and will only be able to access the master data that is appropriate to that role.

Here are a number of other capabilities to be considered:

Consolidate

- Capture attributes and relationships for people, households, and organizations in a rich and extensible data model
- Create a universal ID for each customer and build a cross-reference to each connected system
- Take advantage of an import workbench and UI that helps data stewards manage source system mappings, data loads, and customer profiles
- Track the source of customer attributes and view historical customer profiles
- Integrate using prebuilt, standards-compliant business services and adapters
- Create a single record from multiple sources using configurable attribute "'survivorship" rules

Cleanse

- Leading embedded data-quality tools and integration to third-party data sources
- Cleanse addresses via open interfaces to address cleansing providers
- Identify duplicates using fuzzy matching techniques
- Intelligent merge, auto-merge, and unmerge capabilities to eliminate duplicates
- Enrich the customer profile using pre-built integration with external providers such as D&B
- Supports knowledge-based MDM via out-of-the-box integration with Acxiom
- Includes customer data lifecycle management workflows

Govern

- Manage privacy and regulatory policies using prebuilt rules and events (including Opt-Out preferences)
- Respond to customer lifecycle events with best-in-class workflow capabilities
- Full history and audit trails for rock-solid security and proof of compliance
- Full access to rules engines to support changes requested by governance committees
- Roll back to a prior point in time to undo events such as an unwanted customer merge

Synchronize (Share)

- Provide the “golden record” and selected attributes to all applications and analytical systems
- Provides commonly used functions as business services and Web services
- Supports batch and publish/subscribe to maintain consistent data in real time
- Utilizes any leading middleware provider
- Prebuilt integration to leading front office and back office applications
- Prebuilt integration to leading analytical and business intelligence applications

6. Summary

The trend in today's companies towards a more frictionless, efficient, electronic environment and the movement towards "straight through processing" will only work if it rests on a strong data quality foundation.

But more importantly, without big changes, a lot of companies will continue to struggle. There is relentless pressure on costs, coupled with fairly slow growth in demand. There's no room to lose and then replace existing customers.

Most companies are juggling many different transactional and analytic applications across their corporate headquarters and their global, regional and local operations. At the same time, customers are demanding faster and more complex responses from organizations.

This is all leading to a disconnect that threatens organizations' ability to make decisions and act quickly within their markets.

A UK company, TUI Travel, restated its 2009 financial results due to issues arising from using two separate IT systems after the 2007 merger of two companies, in the amount of $191.8 million US dollars (£117m). This type of persistent operating silo is all too common in the companies I see, but this is probably a worst case scenario.

But look for those costs that MDM, data governance and data quality can help your enterprise avoid, the revenue increases they can help you realize, and the compliance improvement and agility increases they can bring to your company.

Those reductions, improvements and increases are very real. I've worked with enough companies now to know firsthand that MDM is not a fad or a trend, but a real force in the marketplace. And it's experiencing growth of 18% at a time when the enterprise software market as a whole is only growing at about 5.6%.

These are interesting times that we live in. Companies need every edge to compete – the old days of being able to be complacent and to tolerate widespread process and data inefficiencies are gone. Today, businesses need to have the Single View of their master data, and to be able to "Create Once, Use Many" in order to be lean, and indeed, to survive.

The silos that are still so prevalent do nothing but increase costs, hurt business and IT agility, and result in bad decisions being made (because of low quality, inaccurate, inconsistent data).

If we implement Enterprise Data Quality tools, master data management solutions, and remember that these technologies must be driven by business-led data governance organizations, we'll be able to break down those silos, and have enterprise-wide master data repositories that lead to streamlined processes, lower costs, better decisions, and more agile organizations that can compete better in global markets.

Ultimately, that's the vision of master data management and data governance that I try to help my clients envision for their companies. And I've seen enough of them realize it successfully that I know it's realistic, achievable, and provides an outstanding return on investment. A well thought out strategy plus focused delivery with strong attention to detail will succeed in the end.

Oracle Master Data Management: Complete, Integrated, Best-of-Breed

Oracle's suite of products for enterprise master data management is the most complete Master Data Management (MDM) offering in the industry. MDM applications manage customer, supplier, product, and financial data with data governance services and supporting world-class integration and BI components. What's more, Oracle MDM includes pre-built support for both business process, services and data integration to create MDM-aware applications that help better integrate master data to Oracle Applications as well as custom and 3rd party business applications.

For each major master data dimension, Oracle delivers best-of-breed MDM capabilities in the form of an MDM Hub. These hubs provide: data models for the dimension; data steward workbenches to manage the consolidation of the data from all connected systems; data quality tools tailored to the nature of the master data; data governance tools for the business user; extensive application integration to share the governed data with all appropriate operational applications; and data maps for key analytical systems such as data warehouses and operational business intelligence dashboards.

The following hubs are available today along with various industry specific versions:

- Oracle Customer Hub
- Oracle Supplier Hub
- Oracle Product Hub
- Oracle Site Hub

In addition, Oracle MDM includes two important elements for Enterprise Data Quality and Data Relationship Management (DRM).

Oracle Enterprise Data Quality helps companies to achieve maximum value from their business-critical applications by delivering fit-for-purpose data – enabling individuals and collaborative teams to quickly and easily identify and resolve any problems in underlying data. Oracle Enterprise Data Quality is a complete and integrated suite of data quality tools that provide an end-to-end solution to measure, improve, and manage the quality of data from any domain, including customer and product data. With Oracle Enterprise Data Quality products, customers can identify new opportunities, improve operational efficiency, and more effectively comply with industry or governmental regulation.

Oracle Hyperion Data Relationship Management enables your organization to proactively manage changes in master data across operational, analytical and enterprise performance management silos. Business users may make changes in their departmental perspectives while ensuring conformance to enterprise standards. Whether processing financial information such as cost centers, accounts, and legal entities or analytical information such as members, dimensions, hierarchies, and their attributes, Oracle Hyperion Data Relationship Management delivers timely, accurate and consistent master data to drive ongoing operational execution, business intelligence and performance management.

The results for companies who implement Oracle MDM solutions are dramatic. Over 1000 companies and organizations are managing billions of master data records with Oracle MDM. Companies such as Cisco, GE, Fidelity, Motorola, Dell, Symantec, Zebra, LG Telecom, Korean Air, Home Depot, Toyota, Starbucks, Credit Suisse, and Scottrade to name a few are realizing the promise of consolidated, clean, consistent master data feeding their operational and analytical systems.

To learn more about Oracle's complete and integrated MDM offerings, go to: www.oracle.com/goto/MDM.

About The Author

Dan Power (http://twitter.com/dan_power) is the founder and president of Hub Designs. His role there is a combination of best practice expert, industry analyst, client advisor, and thought leader. He's responsible for client strategy and delivery in the areas of MDM, data governance, content marketing and social media.

Dan is the author of more than 30 articles and white papers on MDM and data governance, and is a featured speaker at technology conferences, webinars and vendor events. He writes a column for *Information Management* magazine and is the editor of *Hub Designs Magazine* (http://hubdesignsmagazine.com), a widely read online magazine covering MDM and data governance.

About Hub Solution Designs, Inc.

Hub Designs, a global management and technology consulting firm, specializes in developing and executing high value master data management and data governance strategies.

Through recognized thought leadership, an excellent reputation and a strategic network of partnerships, the firm delivers successful projects to Fortune 1000 clients, who are its best references. For more information, visit www.hubdesigns.com.

About Oracle MDM

Oracle's enterprise master data management (MDM) suite of products consolidates and maintains complete, accurate, and authoritative master data across the enterprise and distributes this master information to all operational and analytical applications as a shared service.

Oracle's suite of products for enterprise master data management is the most complete MDM offering on the market. MDM applications manage customer, supplier, product, and financial data with data governance services and supporting world-class integration and BI components.

For more information, visit www.oracle.com/goto/MDM.

About Oracle

Oracle engineers hardware and software to work together in the cloud and in your data center. For more information about Oracle (NASDAQ:ORCL), visit www.oracle.com.

Oracle Corporation
500 Oracle Parkway, Redwood Shores CA 94065 USA
General inquiries: +1.650.506.7000 or +1.800.392.2999
International: +1.650.506.7000 or +1.800.ORACLE1
www.oracle.com

Hub Designs
188 Whiting Street, Suite 6A, Hingham MA 02043 USA
+1.781.749.8910
info@hubdesigns.com
www.hubdesigns.com | http://hubdesignsmagazine.com