The Information Difference

# Building a Robust Business Case for High Quality Master Data

An Information Difference White Paper

Andy Hayler

February 2010

Sponsored by




# TABLE OF CONTENTS

## EXECUTIVE SUMMARY

Projects addressing data quality or master data management (MDM) frequently struggle to get approved by senior management. Indeed, in our research we found that the top two barriers to data quality projects were the (related) items "it's very difficult to present a business case" and "management does not see this as an imperative". Only around 60% of MDM projects proceed with a proper business case. This is an unsatisfactory state of affairs. Quite apart from being poor practice, if a major project proceeds without a clear idea of where benefits are to come from, then it is at a high risk of cancellation (or outright failure through lack of correct focus) when times become tough and a hard look is taken at all discretionary projects.

There is no reason for any project of this type not to deliver a proper, quantified business case. In this document we explain the key elements of a financial case, preferring financial measures that take into account the time value of money over looser measures which are in common use—especially in the IT press—but are frowned upon by many corporate finance departments. Several well-established project-estimating techniques exist, and data is now beginning to come through on the typical costs of real-life MDM projects, both initial and maintenance costs. The tougher side of the picture is usually developing the benefit case. We draw heavily on both the personal experience of the author and a number of studies to go through numerous real-life examples of the benefits that have accrued through successful data quality and MDM projects, which in some cases come to hundreds of millions of dollars annually.

We recommend a scenario approach to the financial aspect of a business case, demonstrating the robustness of the project to a range of scenarios, not just a single hoped-for set of benefits, since our experience has shown that the benefits that transpire from one of these projects are rarely what are predicted at the outset. We also recommend that post-implementation reviews of completed data quality and MDM projects are carried out, both to draw out useful lessons that other projects can learn from, and to instill a culture of keeping the benefits estimates honest.

Armed with the material in this paper, you will be in a good position to deliver a high quality business case for your data quality or MDM project.

## BACKGROUND

In recent years master data management has become one of the fastest growing elements of the enterprise software market. Data quality initiatives have been around for much longer, but increasingly the two are becoming linked. While a particular data quality initiative could be considered without necessarily addressing enterprise-wide master data, the reverse is not true; every master data management project has a significant data quality component. In one study[1] we found that, on average, 30% of an MDM project was dedicated to data quality. There is a common link between MDM and data quality projects: building a business case for them is problematic. In a

[1] "MDM projects in Practice" December 2009, The Information Difference www.informationdifference.com

separate survey[2], the two top barriers cited by survey respondents were the linked items "it's very difficult to present a business case" and "management does not see this as an imperative". In one survey[3] we found that, on average, just 60% of MDM projects actually implemented had a formal business case.

As we shall see, there are particular aspects of MDM and data quality projects which can complicate the process of building a business case, but we believe that such obstacles can be overcome. In discussions with many companies, we have found a surprising lack of familiarity with the mechanics of how to present a business case, quite apart from any difficulties in estimating actual costs and benefits. Since an MDM project in particular tends to be quite large (an eight-person project team is average according to one of our surveys, and we have encountered projects with teams of over 100) it may seem surprising that such significant projects get off the ground without a formal business case. Anecdotally, it seems as if many are backed by an influential sponsor, and we are often told, "We don't need a business case since the project is backed by the CFO," (or some other executive). This is, in our view, a dangerous approach. An enterprise-wide MDM initiative is likely to be a multi-year project, and executives in major corporations move around at very regular intervals, so a project dependent on one influential sponsor is highly exposed when that sponsor moves jobs, or when some external shock, such as a recession, causes all discretionary projects to be reviewed.

Consequently, we will go into some detail in this paper about the basic elements of a business case before discussing the peculiarities of data quality and MDM projects, and draw on some real-life examples in order to illustrate how to present a high quality business case.

## THE COMPONENTS OF A BUSINESS CASE

A good business case explains the justification for a project, and describes in *quantified terms* why that project is a sensible investment when stacked up against other projects which a company could undertake with those resources and money. Every company has a finite amount of capital and people's time, and there are usually a lot more ideas for how to spend money than there is money available. A well-run company will have a disciplined approach to allocating money and resources, and will stack up and compare alternative projects before deciding which ones to carry out. Since most potential projects will be different from one another, some objective mechanism is needed to compare their effectiveness.

In some situations a project is mandatory, e.g., in response to a regulatory change or emergency, but in most cases there is always a "do nothing" alternative, i.e., carry on with the status quo. In order to prioritize projects we need to stack them up in a way that is directly comparable. In other words, we need to provide an objective way of choosing project A over project B rather than just listening to who makes the most eloquent case or shouts loudest.

The way most large enterprises go about selecting investment projects is to compare a number of financial ways of assessing the project. Generically, this is often referred to as the "return on

[2] "The State of Data Quality" May 2009, The Information Difference www.informatoindifference.com
[3] "MDM Projects in Practice", December 2009, The Information Difference www.informatoindifference.com

investment" (ROI) of the project, but as we shall see there are certain specific ways of measuring this that are better than others. So, what are these measures?

To begin with, ROI itself can be defined as:

**ROI = (gain – cost)/cost**

If a project costs $1million, but returns $1.5 million in benefits, then its ROI is 0.5, i.e., 50%. However, this is a very crude measure, since it does not take into account the time period over which the costs and benefits are spread out. In real life, a project may take years to deliver, and the benefits usually come after the project has been completed. It is important to realize that a benefit projected for this coming year is actually a more valuable commodity than one projected three years out, since there is considerably less certainty attached to whether a benefit will really occur at some time in the future, given the realities of the business world.

The old saying, "A bird in the hand is worth two in the bush," is relevant here. After all, if someone offered you a choice of a gift of $1,000 today, or $1,000 in three years' time, you would obviously choose the money right now, rather than taking a chance on it turning up in three years' time. Moreover, if you take the money now then you can invest it and that $1,000 has the chance of turning into even more money. So even if you exclude uncertainty, money now is worth more than money in the future since you can earn interest on money that you receive now.

Consider a second question. Would you rather have $1,000 now or $1,100 in three years time? That is a trickier question to answer. Even if you exclude the possibility that some unexpected event may occur that could prevent you getting the money at all in the future (such as, if the person or company offering you the money were to go bankrupt), then whether it is better to take the $1,000 now or $1,100 in three years depends on the amount of interest you can earn on that money, and with what degree of certainty you have of achieving that rate of interest. For example, suppose you could be pretty sure of getting an interest rate of 4%, perhaps via a government bond or some guaranteed savings rate from a reputable bank that has negligible chance of going bust (not such an easy thing to find these days). In that case, your $1,000 invested today at 4% will earn you $1,124.86 in three years, due to compound interest:

**Initial investment: $1,000**
**Interest rate: 4%**

**Value at end of year 1: $1040 (4% interest on the $1,000)**
**Value at end of year 2: $1081.60 (applying 4% interest to the accumulated $1,040)**
**Value at end of year 3: $1124.86 (applying 4% to the accumulated $1081.60).**

In other words, you'd still be better off taking the $1,000 today and investing it, than you would be if you waited for an inferior $1,100 in three years. However, if you could only get an interest rate of 3%, then your money invested over three years would only be worth $1,092.73, so you'd be a little better off taking the promise of $1,100 in three years, *provided* you could be pretty certain that this promise will be kept. If you could only get 2% interest then your money would be worth $1061.21 in three years, significantly less attractive than $1,100 in three years' time.

Because of this importance of the effect of time on the value of money, most businesses do not use crude ROI as their main assessment of the project investment quality. Instead, they use measures which work on the concept of discounted cash flow. This may sound rather esoteric but in fact the idea behind it is simple. In order to take account of the fact that money now is better than money in the future, a particular interest rate, called the *discount* rate, is chosen.

What happens is that the future expected returns are reduced by the discount rate, applied to each year as needed, on a compounding basis just as with the example above. Essentially the discount rate is like a reverse interest rate; instead of making your money grow over time into the future, the discount rate starts in the future and works backward. It can be seen in our example that $1,000 today is worth a bit more than $1,100 (actually $1,124.86) in three years at a 4% interest rate, but a bit less ($1092.73) at a 3% interest rate.

Now, actually calculating the rate needed for today's $1,000 to become $1,100 in three years' time is a little fiddly to work out (clearly it is somewhere between 3% and 4%), but in these days of spreadsheets we can take advantage of a built-in Excel function (called "IRR" in this case) to work this out for us. It turns out that today's $1,000 will turn into $1,100 in three years at an interest rate of 3.23%. This rate is called *the internal rate of return* (IRR).

Let's take another example to clarify this further. Suppose you bought a house for $100k and figured that you would be able to sell it in three years for $130k. In this case, it turns out that the internal rate of return (IRR) of this investment idea is 9.14%. This would be a very healthy return (good luck finding a deposit account that will pay 9% these days). However, predicting house prices in the future is an uncertain business, so it may be more realistic to apply a *risk premium* to the calculation. Deciding level of risk is subjective, but suppose we chose 5%. In that case, the risk-adjusted return would be 9.14% - 5%, i.e., just over 4%. This is still an acceptable return. Indeed, any positive value here may be deemed to be acceptable. However, it can quickly be seen that the strength of this investment idea is very susceptible to the assumptions made. Suppose our house is more likely to sell in three years for just $115k. We will still have made a nominal profit (with a 4.77% IRR), but our risk free return (i.e., allowing for the 5% risk premium) is 4.77 – 5, i.e., slightly negative. We might decide that this is just too risky to proceed.

It is also important to understand that the more distant in time our expected benefit, the worse the IRR will be. If we expect to sell for $130k in five years instead of three years, then our IRR is 5.39% instead of 9.14%. This is because the discount rate is applied over a longer period. In other words, a specific amount of money five years from now is riskier than the same amount of money three years from now.

So, we now understand IRR. It is effectively the interest rate that you earn from your investment. There is one more measure that is important to understand, and that is *net present value* (NPV). Returning to our house-buying example, our IRR was 9.14%, which is fine, but we might like to know what the actual monetary return would be, adjusted back to today's money. We expect to get $30,000 in profit, but that is in three year's time—what would that amount actually be worth today? This clearly depends on the interest rate (*discount rate*) that we are going to use. Let's take 4%. We could calculate this out by hand, but to save time using the Excel NPV function, we discover that the NPV here is $14,970.70. It is noticeable how much less this is in absolute terms than the $30,000. If we use a higher discount rate, then the NPV will decrease. At a 5% discount rate the NPV becomes

$11,713.23. The more we ratchet up the discount rate, the worse the NPV. At an 8% discount rate the NP is as low as $2961.29, and at 10% it is actually negative -$2117.34. Any positive value for an NPV indicates a reasonable project, but if the NPV is negative then we should generally avoid it.

How high a rate we use is determined by our attitude to risk. The more confident we are about the house price several years into the future, the more comfortable we will feel with a lower discount rate. Out of curiosity, at what discount rate is the NPV exactly zero? This is our old friend, the IRR, i.e., 9.14%.

So, what does all this accounting jargon have to do with a business case for an MDM or data quality project? NPV and IRR are in fact the two measures which most corporate finance departments use to determine whether or not to approve an investment project. As a minimum, the NPV should be positive using a conservative discount rate. The actual rate chosen will vary from company to company, but usually represents the costs of raising capital of that company (whether debt or equity). It is beyond the scope of this paper to discuss the subtleties of costs of capital, but essentially you will find that it is normally a number in the 15%-20% range. Let's say our company uses 15% for the purposes of this paper (this is sometimes called the "hurdle rate"). As long as the NPV is positive using the corporate hurdle rate, then the project should in principle be carried out from a pure accounting viewpoint. In reality, there are finite resources and competing projects, so the projects with the higher IRR will usually win the battle. A project with a 50% IRR should always be selected over one with an IRR of 25% (at least in purely financial terms), since this will deliver a much better financial return for the company and its shareholders.

So, we now understand IRR and NPV, and are almost ready to look at the specifics of business cases for MDM and data quality projects. Just before we do, I will briefly mention one other commonly used measure of an investment project, the *payback period*. This is simply how long a project takes to return the original investment. Hence, a project costing $1 million to deliver with a $500k annual benefit has a two year payback period. The shorter the payback period, the better the project is. Payback period is easy to understand but is actually a crude measure, since it does not take into account the time value of money properly. This is why NPV and IRR are generally preferred for assessing investment cases. (There are some more esoteric measures, but these are all we are likely to need in most companies.)

So, armed with our financial measures, we can start to look into how to apply these to the problem in hand: presenting a business case for a data quality or MDM project. For this we now know that we need to estimate the initial project costs, how long the project will take, and the benefits that the project will bring. As is usual with business cases, we will discover that estimating the costs and elapsed time is easier than estimating the benefits.

## COSTS

Although IT projects are notoriously tricky to successfully estimate, at least there are some reasonably well understood techniques for estimating IT projects. It is not within the scope of this paper to discuss IT project estimating in detail, as this has been the subject of numerous books. However, suffice it to say that every self-respecting IT project methodology has one or more techniques to help estimate the effort needed to complete a project. One way is to split the project

up into small tasks (software installation, design, testing, etc.), each of which is well understood enough to be accurately estimated. This assumes that you actually do have a good grasp of all the elements necessary for an MDM or data quality project, but in essence these will have the same broad components as any other: feasibility, technology selection, logical design, physical design, implementation, testing and going into production. There are other general estimating techniques, such as the estimation of "function points" to be delivered, and there are some good books which have covered such techniques in detail (see Further Reading).

Let us look, though, at some specifics of the type of projects in which we are interested. In 2009, The Information Difference conducted the survey "MDM Projects in Practice". In this survey, the respondents indicated that, on average, their MDM project involved a project team of eight and took six months, involving the management of an average of 3 million master data records. Clearly, these are just averages, but these are the numbers which we will use to build a worked example business case, since they are at least real-life figures. It is fair to say that these projects are mostly implementations of a single master data hub covering a major data domain, e.g., "customer data in Europe" rather than being truly enterprise-wide MDM projects. We have worked on MDM projects on a global basis which cost over $100 million, to give a sense of the potential scale of such projects.

In order to complete the cost estimates, we need to come up with a "fully-overheaded" cost of a staff-day. Companies usually have such figures easily available from their finance departments. It is important not to count just average staff salary, but also benefits and all the inherent overheads a company has in addition to its wage bill (office rent, accounting fees, equipment costs, etc.). Of course the actual number will vary by company and by country, but let's, for the sake of argument, assume that this figure is $600 per day. In practice, we may need to adjust this to allow for the costs of external consultants, who are usually more expensive than internal staff. What proportions of internal vs. external staff are typical? In our 2009 survey we found that the averages were:

**Internal IT staff 25%**
**Business staff 40%**
**External consultants 35%**

If we allow an average of $1,000 for the external consultants, then we can come up with a blended cost of a person day of $740.

For our average project of 48 person months, assuming 20 working days per month, this means that we have staff costs of:

**48 * 20 * 740 = $710,400**

To this we also have to add software costs and any hardware costs associated with the project (let's say we need some servers costing $20k). Licenses for MDM software (assuming that the project is using a package rather than being custom built; around 80% of MDM projects use a package according to our figures) vary significantly by vendor, with various pricing models in use. Assuming we have a conventional perpetual license, let us for the purposes of our example say that we have:

**License costs: $200k**
**Annual maintenance: $40k.**

In that case we can now see that we have our project costs:

**People costs:** **$710k**
**Software costs:** **$200k**
**Hardware cost:** **$20k**
**Total:** **$930k**

How about the ongoing annual costs? Again we can turn to some real data. In our 2009 survey we found that, on average, MDM projects cost 20% of their implementation costs. (This was based on projects that had completed. Interestingly, those who were just planning projects estimated less than this.) We can therefore assume:

**Ongoing people costs:** **$142k**
**Ongoing software maintenance:** **$40k**
**Total:** **$182k**

This gives us the cost side of the picture for our fictitious MDM project. We will shortly turn to the benefits.

The Information Difference has a specific MDM project estimating tool, which is calibrated to a base of real-life MDM projects and, based on a number of characteristics of the project, will provide a ballpark estimate. If you are interested in learning more about this tool, please contact us directly.

# BENEFITS

## Issues

While it is no piece of cake to successfully estimate any IT project costs, it is generally tougher to estimate benefits. Apart from anything else, there are no generic estimating techniques to draw on, since the benefits associated with a project are specific to the organization: the benefits for a pharmaceutical company of high quality master data will be quite different to those a bank would find, and different again to those a retailer would identify.

There is a second problem. Even if it is clear that delivering better and more consistent master data is a good thing, it is difficult to separate out the benefits that are purely associated with the better data, as distinct from other improvements. For example, if a data quality project is implemented that improves the quality of customer data, and customer satisfaction as measured by an annual survey improves by 10%, just how much of this improvement can be attributed to the data quality project? Doubtless there were other projects running at about the same time that also aimed to improve customer satisfaction, and each of these might reasonably claim to have had an impact on customer satisfaction.

The impact of improved quality master data is frequently indirect. It may be that, after a successful MDM project, the number of sources of product data in a company was reduced from nine to just one, and following on from this, a new e-commerce initiative went better than expected. Some of this was down to the ease with which the e-commerce project was able to access good quality, consistent product data, but to what extent can you really attribute the benefits to the MDM project? This is a common problem with IT infrastructure projects which deliver greater flexibility

for other projects that follow. It is hard to be sure how much of the later benefits are due to the better infrastructure, as we are not in a laboratory where we can run a double-blind trial to compare the effect of the project against a controlled situation where no project was undertaken. As noted earlier, in some cases a project may go ahead due to some regulatory reason without a business case, but we will concern ourselves here about projects that need to be justified financially.

Despite these issues, there are some categories of benefits that we can examine.

## Data Quality Projects

If we first look at projects that address a specific data quality issue (say, the quality of supplier names and addresses) but do not try to provide a reconciled set of master data, then it is likely that most of the benefits will focus around operational processes. Examples include:

- Reduced marketing costs due to cleaner prospect data. Sending direct mail out by post costs money, and if you have a significant degree of duplication in your mailing lists then cleaning these up will reduce your postal bill. (This is quite apart from improving your image. Who has not received two or three identical envelopes through the post with some special offer? This hardly inspires consumer confidence in the company sending the duplicate mail.)
- Duplicate customer accounts can easily be set up by accident unless you employ data quality software, due to variations in people's names (e.g., A. Hayler, Andy Hayler, Andrew Hayler) as well as common misspellings.
- Poor quality customer data can have unexpected consequences. One UK bank which we worked with discovered 8,000 customers that were—according to their systems—over 150 years old. This was merely mildly amusing until their marketing department decided to try and cross-sell life insurance to current account holders. Imagine the kind of actuarial quote you would receive if you appeared already to be 150 years old.
- In a B2B world, the same issues arise. The author worked on a project in Shell Australia where it transpired that as many as 80% of all the commercial customers listed in the customer database were duplicates.
- Poor quality customer data can result in missed deliveries and late and challenged invoices (which has an effect on cash flow).
- Incorrect spatial data led to an exploration well being drilled directly into an existing production well in the 1990s in the North Sea. It was only because the producing well happened to be down for maintenance that a major environmental incident was averted, and instead resulted in an engineering problem costing "only" several million pounds to fix.

The above are just examples, but hopefully will give you some idea of where to look.

In the case of product data rather than customer data, other benefits categories often occur. Companies often manually go through product files in order to remove duplicates (companies frequently report that 15-30% of their product catalogue entries are really duplicates) and accurately classify products. This process requires some domain expertise and, because it is manual, is inherently unreliable since it depends on the skills of the individual. Alternatively, companies outsource this process, and indeed there is a mini-industry in India of combing through product files on behalf of western companies. It is clear that using automated tools to deal with product quality is a better approach, but it may be difficult to define the rules that are needed to categorize products properly and identify duplicates, since product files are frequently unstructured and can have

hundreds of attributes. However, if data quality improvements can be successfully made, then a number of categories of benefits can occur:

- In the case of getting retail product catalogues online for e-commerce, bringing on board a new product and ensuring it is properly categorized for a website can be a tedious process. For example, in one project in Home Depot Canada, it initially took as much as ten weeks for each new product to be fully on-boarded for online sales, but after a major clean-up of product data, implementation of a single product hub and streamlined processes this on-boarding time was slashed to ten days; with thousands of products involved this was a significant savings.
- The key here is in estimating the time taken to do the task manually, and comparing this with the situation after an improvement project. For example, one medium-sized manufacturing company that we spoke to had a catalogue of 90,000 products with 15% duplicates, and found that the manual effort in correctly classifying the products was costing them over $85k annually.
- If you can get a proper view of your current product inventory, then this can have a number of categories of savings. You can avoid wasted purchases by having visibility across all of your sites, improve stock management, sell obsolete items, reduce storage costs and in some cases reduce the downtime of operational equipment by more efficiently sourcing parts for its operation.

Clearly, the specific benefits around product data depend upon the size of your product catalogue, the state of your data and the particular efficiencies that can be achieved by improving product data quality. However, again you can see from the above examples some of the likely categories of benefits.

## MDM Projects

An MDM project, by definition, will improve the consistency of master data across at least part of an organization. Hence, we should be able to look at categories of benefits that are relevant to this:

- Time spent reconciling multiple sets of operational data today
- Better business intelligence, hence (hopefully) better business decisions
- Lower IT costs due to reduced numbers of interfaces
- Improved customer satisfaction
- Greater strategic flexibility.

In one survey[4] we discovered that:

- 37% of large companies do not calculate profit margins consistently across divisions.
- 38% cannot quickly calculate spending by suppliers across divisions
- 23% struggle to determine gross sales to global customers.

This kind of problem is more than just irritating. One company in which the author worked lost a huge global account because the customer became frustrated that they were unable to be dealt with on a consistent global basis, a decision that cost tens of millions of dollars.

[4] MDM Adoption Survey, March 2008, The Information Difference

In another example, one global company was able, following a project that delivered consistent sales and profitability data across Europe (previously each country subsidiary had inconsistent profit margin definitions) to gain a clear picture of its profitability by country, product and major customer. It discovered that in one subsidiary there was a large number of barely profitable or even unprofitable transactions with commercial customers involving one of its premium brands. A decision was taken to significantly raise the price of the product overnight, which caused as many as 20% of the affected customers to leave but had a net $25 million dollar margin improvement to the corporation (the customers that left were largely the unprofitable ones).

In general, it will be easier to quantify time spent on reconciling data, and on looking at the costs of current interfaces, than on things like better BI, customer satisfaction or flexibility. The latter may be very real, and even more important than the operational costs, but are hard to predict. (For example, the $25 million improvement example above was entirely unexpected by the company, who had no idea that they had this problem before the project was completed.)

Although the examples above are just that and cannot be directly be applied to your particular MDM project, it can be seen that these types of benefits are ones that you can look for in your particular situation, and seek out by interviewing the relevant business staff, who will usually have an excellent idea of the opportunities and issues that today's situation is causing.

## Working Backwards

A useful trick if you are having trouble estimating benefits is to start with known problems and work backwards from the costs, which are usually easier for people to identify and estimate. For example, if it is known that 20% of invoices are incorrect due to poor master data, then it should be possible to identify the costs of handling each invoice that is challenged. Suppose a company issues 5,000 invoices a year (of average value $50k), of which a fifth are challenged, each of which results in an average ten-day delay while the situation is resolved, costing on average four hours of person time (half a day) effort to resolve. Assuming a person-day costs $600, then that is 5,000 * 20% * 0.5 * 600 = $300k.

Moreover, we also have a ten-day delay in money reaching the company accounts, where money can be earning interest. If a 5% annual interest rate could be obtained on such money, then this delay to a fifth of our invoices affects $50 million of income, costing a further $68k of lost interest. These numbers are of course made up, but they give you an idea of the kind of process that can be gone through. Of course the MDM project is not going to magically resolve every invoice problem, but even if it fixed half of the issues then it can be seen that there would be an annual benefit of $150k+$34k =$184k, which will certainly help the business case.

## Risk

Risk to a corporation can take many forms, and it would be impossible to assess every risk, whether competitive, regulatory or market-related. Some risks could have serious consequences but be quite unlikely, and the financial impacts of risks may be very hard to assess. Nonetheless, poor quality master data can have a bearing on some important risks, and not merely regulatory ones. (In some industries, e.g., finance, you may be required to measure certain aspects of data quality, or risk a fine.)

Poor master data that can have serious consequences, for example, in the world of financial trading. Taking a case the author is personally familiar with, oil traders use financial instruments called derivatives to bet on the future oil price, either to hedge supply risk or for speculative purposes. One risk that needs to be considered is that the company that you are trading with (the *counterparty*) goes bust; clearly you want to know your maximum exposure on a given day to a default of a counterparty. One major global oil company thought that it had a reasonable idea of its exposure to Enron, until Enron actually folded in December 2001. It transpired that a whole series of companies with which the company traded ultimately were owned by Enron, meaning that the losses the oil company experienced were several times greater than their risk model had predicted. More recently, the banking default at Lehman Brothers in September 2008 highlights the very real issue here, since it is often far from obvious who the ultimate owner of a company actually is, as there may be a maze of holding companies to navigate in order to be clear as to who is the ultimate owner; if systems do not correctly reflect the ownership structure (which of course can change over time) then exposure can be worse than risk models indicate.

It is obviously difficult to accurately quantify risk of this type in a business case for an MDM project, since the risks that you are (hopefully) mitigating are presumably unlikely to occur. Indeed it would be hard to prove that your project really mitigates such a risk until a disaster unfolds. It may be that you merely note the improved risk in the "soft" benefits of your business case. However, depending on the particular risk, there may be some historical data to draw upon that could enable a quantified estimate to be made.

## Real Life Examples

What real-life examples are there? In talking with and working with a variety of MDM projects around the world, we have encountered a series of very interesting specific costs and benefits. These include:

- Following the rollout of a data warehouse and MDM project across Asia Pacific, it was discovered that a pricing error, previously undetected, had led to a well-known brand being sold for no profit across the region for over a year; the money involved was over $50 million (multi-national company).
- Inconsistent master data led to considerable duplication of marketed products across countries, many of which were literally identical products addressing the same need, but in distinct markets; the removal of this duplication, where four times as many distinct products existed as were needed, easily paid for the $26 million project (energy company).
- One consumer packaged goods firm completed a very large, global project to improve its' master data, taking several years to fully implement. Although the project cost over $100 million, it found $400 million of procurement benefits alone.
- "Poor data quality and consistency has led to the orphaning of $32 million of stock just sitting in the warehouse that can't be sold since it's lost in the system" (manufacturing company interviewed in an Information Difference Survey in 2009).

Of course these examples are not necessarily directly relevant to your particular situation, but they do give a sense of the scale of what has been achieved elsewhere, and give some idea as to the types of benefits that can occur in practice.

## Scenarios

The example we saw earlier showed just how sensitive a business case is to the level of claimed benefits, which are inherently estimates and cannot be certain. In order to make the business case more robust, a useful technique we have used with clients is to develop a range of *scenarios* for the business case.

For example, suppose we have identified annual benefits of $600k. Perhaps this was based on cost savings from better operational processes, which had currently been assessed by the business as having avoidable errors costing $1million annually. It was felt that the project, if delivered successfully, would fix 60% of these problems, so the benefits are estimated at $600k annually.

However, how accurate is that 70% (????) estimate? It may be more prudent to run the business case with a range of estimates of improvement, such as:

- Pessimistic 50% improvement, i.e., $500k saving
- Target 60% improvement, i.e., $600k saving
- Optimistic 70% improvement, i.e., $700k saving.

Is this business case solid through the range of estimates? It may turn out that the business case works at the target level of benefits, but fails at the pessimistic one, for example. It is prudent to run a range of scenarios like this. Benefits rarely turn out exactly as expected (sometimes better, sometimes worse) so it is sensible to see how the project will perform under a range of assumptions.

## WORKED EXAMPLE

We have now, in some detail, been through how to estimate a set of costs and benefits for a sample MDM project. Let us now pull this information together using the numbers from earlier. We will assume that the project takes six months to complete (so we should allow for half the annual maintenance costs in year one, and also half the annual benefits). We are now able to produce our business case:

| Year | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Project costs | 930,400 | | | |
| Maintenance | 91,040 | 182,080 | 182,080 | 182,080 |
| Benefits | 300,000 | 600,000 | 600,000 | 600,000 |
| Cash flow | -721,440 | 417,920 | 417,920 | 417,920 |
| Discount rate | 15% | | | |
| IRR | 34% | | | |
| NPV | $202,405 | | | |

We see that our net present value is a positive number, so the project is at least inherently sound. We have an IRR of 34%, which is respectable, and crucially is greater than the company hurdle rate of 15%. Consequently, this project should squeeze past a finance review. Whether it actually made it past scrutiny, however, would depend on the quality of other competing project proposals (and the various softer real-life issues, like how well politically connected is the project sponsor).

We can see just how sensitive the business case is to the benefits estimate. In this case, if the benefits estimate had been just a little less at $500k then the NPV would have been negative (-$39,615) and so would, at this level of benefit, not be sound. In fact, it is fair to say that this particular project business case is somewhat marginal, since any benefits estimate is unlikely to be very precise. If the benefits were $700k then the IRR would shoot up to 57%, which is a highly respectable return for a project. However, the fact that the business case fails in the "conservative" scenario makes it somewhat vulnerable to review.

Of course, a complete business case would also contain textual background and justification for the project, explaining where the benefits came from, and any "soft" non-quantified benefits that were expected from the project. However, at its core, a good quality business case should have hard numbers, quantifying the net present value and internal rate of the return of the project. A well-argued business case with these numbers at its heart is likely to gain backing from upper management.

## POST-IMPLEMENTATION REVIEW

It is not enough to just make a claim for a series of benefits. In a well-run company there will be a review after the project has been running for some time, in order to see whether those benefits actually transpired. In reality, the actual effects of a project will probably be different from those predicted. Some claimed benefits will not have transpired or been less than expected, while other unexpected benefits may well have occurred. By documenting these lessons, you can improve the quality of future business cases and potentially see opportunities for further improvement projects.

The 2009 Information Difference Study "MDM Projects in Practice" found that 40% of companies that had completed an MDM project actually carried out a post-implementation review. In our view, it is important that such reviews be carried out and become part of the corporate culture in order for improvements in the quality of business cases to be made. After all, if you do not carry out such a review, there will be a temptation to make over-optimistic benefits claims in order to justify favored projects.

A post-implementation review should include:

- A record of how the project went compared to its original objectives, budget and benefits
- A record of the achievements made by the project
- Lessons learned from the project, which can be applied to future projects.

It is important to understand that post-implementation reviews should not be aimed at apportioning blame for mistakes, but to honestly assess what went well, and less well, on a project in order that lessons can be learned and shared for future projects. If a "blame culture" develops, then these reviews will fail, as people cover their tracks and lose trust in the process.

## CASE STUDY

In 2009, a large pharmaceutical company undertook a project in its UK production procurement area. Traditionally, the various engineering depot sites had been independently responsible for managing engineering inventory such as spare parts. This organizational structure, and limitations in the search capabilities of the ERP system, meant that it was difficult for staff to find whether a particular spare part was already in stock (from procurement perspective, it was almost impossible to obtain an aggregated view across sites of what, how many and where spares were procured from). Frequently, additional stock was purchased which could have been provided from existing inventory from other sites within the network. With over £30 million in inventory at five main sites, this was a significant issue. These five sites dealt with greater than 450 suppliers of engineering spares and had 65,000 inventory items. The items were inconsistently classified, meaning that the company could obtain only very high level information on spending by category. Moreover, as much as 80% of the inventory was "non-moving", suggesting that major savings could be made in inventory. In one extreme case, analysis revealed that the company was holding spares in one area that would last 90 years based on current usage.

A pilot project was conducted which involved extracting the inventory records from the ERP system, cleaning up and properly classifying each record and loading these up into a specialist master data hub. It was decided to leave the maintenance of inventory records within the ERP system at this stage, so an interface was built feeding the clean data back into the ERP system. The objective was to obtain much improved visibility into the engineering inventory, providing much more granular categorization into commodity groups, and so allowing staff to find parts when needed much more easily. The improved categorization would allow better sharing of stock, a reduction in inventory and hence improved working capital. Moreover, it would be possible to rationalize the lengthy supplier list, allowing more strategic sourcing and bringing further advantages due to greater economies of scale and better negotiating leverage with suppliers.

The project business case identified £2.2 million of annual benefits in the first year (through reduction in working capital and sourcing for less); a somewhat conservative scenario. (A higher level of savings was felt to be possible, but this would require a higher level of business effort.) Even given some unexpected internal IT costs, this translated to a net present value of over £3.5 million for the project, with a stellar IRR of over 400%. The project went live in late 2009, and plans are now underway to extend the project to further areas.

## CONCLUSIONS

We have seen how a well-constructed business case should not only discuss in broad terms the reasoning behind the project and the broad benefits that will accrue, but should also deliver a quantified financial case. In particular, if you are putting together such a financial case, you should eschew measures that do not properly take into account the time value of money, such as ROI and payback period. Instead, you should focus on net present value and internal rate of return, which properly account for this, and more to the point are the measures which most well-run companies will seek when assessing competing project proposals. We have seen the typical hurdle rates used in

such project assessments, and so have a sense of what a project needs to deliver in order to be likely to gain backing.

We have examined a number of techniques for estimating project costs, discussed survey data showing average numbers for typical MDM projects, and have learned what real-life MDM projects are experiencing in terms of annual maintenance costs. This gives us one-half of the picture necessary to deliver our business case.

We have also looked at both for data quality and MDM projects, the typical categories of benefits that are likely to occur, and have seen several real-life examples of actual benefits delivered in both kinds of projects. We have learned how working backward from the costs of issues today is an effective technique in fleshing out the benefits that it is reasonable to claim for a data quality or MDM project. We have also learned how it is prudent to run the business case through a range of scenarios, and how it is important to go back and check on what really happened with the project after it went live, by means of a post-implementation review.

Armed with this information, you should now be in a good position to deliver a high quality business case for your data quality or MDM project.

## FURTHER READING

Some good books on software estimating (not specific to MDM or data quality) include:

- Software Engineering Economics, Barry Boehm (ISBN 0138221227)
- Controlling Software Projects, Tom de Marco (ISBN 0131717111)
- Dreaming in Code, Scott Rosenberg (ISBN 1400082463)

## ABOUT THE AUTHOR

Andy Hayler is one of the world's foremost experts on master data management. He started his career at Esso, becoming the youngest ever IT manager for Esso Exploration before moving to Shell as Technology Planning Manager of Shell UK. Andy then became Principal Technology Consultant for Shell International, engaging in significant software evaluation and procurement projects at the enterprise level. He set up a global information management consultancy business which he grew from scratch to 300 staff. Andy was architect of a global master data and data warehouse project for Shell downstream which attained $140M of annual business benefits.

Andy then founded Kalido, which under his leadership was the fastest growing business intelligence vendor in the world in 2001. Andy was the only European named in Red Herring's "Top 10 Innovators of 2002". Kalido was a pioneer in modern data warehousing and master data management.

He is now co-founder and CEO of The Information Difference, a boutique analyst and market research firm, advising corporations, venture capital firms and software companies. Andy is a regular keynote speaker at international conferences on master data management, data governance and data quality.

## ABOUT THE INFORMATION DIFFERENCE

At the Information Difference (www.informationdifference.com) we offer in-depth analysis of the master data management industry. We offer in-depth profiles of the MDM vendors, assessments of the marketplace and white papers discussing key issues and best practice. If you are contemplating an MDM project, we can advise you on strategy, vendor selection and best practice. We carry out primary market research and can help you with MDM project justification and return on investment.