THE VALUATION OF DATA AS AN ASSET: A CONSUMPTION-BASED APPROACH

INTRODUCTION: YOU NEED TO VALUE IT TO OWN IT

There is a case to argue that if companies managed their finances as badly as they manage their data, they would probably be out of business by now. Data used to be confined within firewalls and systems but in the digital age, it is 'happening' everywhere.

In this world of 'big data', having the right data available at the right time is crucial; and it is an imperative that has moved far beyond the remit of IT alone.

Data ownership is a 'business' function; but establishing and motivating already hard-pressed business people to implement best practice around data is difficult - especially where there is no regulatory or internal compliance imperative, unlike the financial services industry. Yet effective data ownership is crucial to good governance.

Gaining hearts and minds might be more achievable if data owners are tasked with managing a cashvalued asset; something that will increase or diminish in value in direct relation to the way it is managed.

Data is a valuable commodity. Most organisations recognise in principle that their data is an asset, but only pay lip-service to the way the concept is applied in practice. Does it ever appear on a balance sheet?

Is it ever assigned a monetary value? Why is it so difficult to value data, in the same way as you would any other corporate asset: machinery, buildings or even employees? Admittedly, data is intangible; it is not a physical entity. And in accounting terms, it does not follow the normal rules. Yet, it has a very demonstrable value in terms of usage and is increasingly being recognised within company valuations - the Facebook IPO in May 2012 demonstrated in the clearest possible way that data can have a very real financial value on the company balance sheet - even if this value was initially overstated!

Companies have a lot to lose from avoiding the issue. If data is not valued as an asset, there is a tendency for it to be seen as an IT expense and, as such, it is under pressure to be reduced, rather than attracting the investment that could improve its value further.

I am putting forward a practical model for the valuation of data as an asset; a model based on the consumption of data. The aim is to make the 'intangible' tangible; to provide the data owners with monetary values that can be managed and improved in a tangible way. Thus, data valuation forming the basis of a key performance indicator (KPI) to motivate each owner to initiate data improvement actions that can be measured and rewarded - ultimately leading to improved confidence in data and the ability to make better decisions.

Once a culture of data valuation and data governance has been established in an enterprise, the extension of the concept to valuing data for balance sheet purposes is a logical further refinement.

This is a practical approach to data ownership and valuation; a 'top down' approach, driven by senior and middle management, involving business managers not just IT; an approach which can be rolled out incrementally across the organisation, across all data repositories.

You need to value data in monetary terms, i.e. as a true asset, to own it.

PART 2: DATA IS DIFFERENT AND NEEDS NEW RULES FOR OWNERSHIP

Like other corporate assets, data can be assigned a cost (i.e. how much does it cost to obtain and maintain); and it has a utility (recognised by where, by whom and how often it is used). It has a number of characteristics, however, which make it different from other corporate assets.

These differences were summarised very well by Daniel Moody and Peter Walsh1, as seven 'laws', in a paper which talks about information - but the same principles can be applied to data.

It was these laws that led me to think about the consumption of data being its primary utilisation and hence an effective measure of its value. They are summarised below, with some of my own comments added:

Law 1: Data is (infinitely) shareable without any loss of value.

Law 2: Its value increases with use (unlike many assets, eg. vehicles, plant and equipment which depreciate in value with use).

Law 3: Data is perishable. Unlike fine wines, the longer you keep it, the less use it is. Unfortunately, many organisations hold on to their data for far longer than required for operational or regulatory needs.

Law 4: The value of data increases with accuracy (We would probably say 'quality' now).

Law 5: Value increases when combined with other data and/or information.

Law 6: More is not necessarily better ('big data' issues bear this out!).

Law 7: Data is not depletable; in fact quite the opposite - the more you use it, the more you have (Glazer, 1993). Conversely, if data is not used, it becomes a liability because despite the cost of storage and maintenance, it is adding no value to the organisation.

Similarly, data that is being replicated many times not only complicates ownership, but often adds cost rather than value to the operation, thanks to the need to maintain separate data storage, develop interfaces or even manually reconcile multiple versions of the 'truth'.

The unique characteristics of data must be recognised by everyone involved in the data ownership process if the data under their care is to be leveraged fully to the benefit of the company. A 'philanthropic' approach should be adopted. Data is not territory to be conquered and fenced off; indeed, the more it is consumed, the more valuable it becomes. The data owner should not be a gatekeeper; there must be no 'keep out - private property' signs and no protectionism, albeit whilst ensuring that all change occurs under the auspices of good data governance and the appropriate security standards.

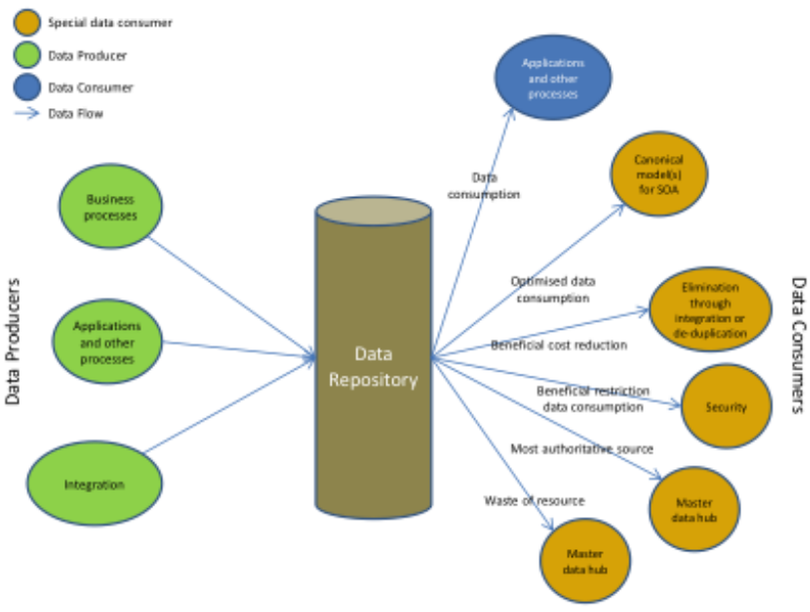
This will require changes to organisational structures and cultural attitudes; something that will evidently not take place overnight.

PART 3: A CONSUMPTION-BASED APPROACH TO DATA VALUATION

My model relies on data ownership at a data repository level (a logical grouping of data, possibly onedatabase or elements of many databases, grouped into a data 'theme' - customer, for example). It takes account of data 'cost' (how much

1 Measuring the Value of Information: an Asset Valuation Approach, presented at European conference on Information Systems, June 1999

does it cost to obtain and maintain); and its utility - where, by whom and how often is it used.
The model works by identifying the 'data producers' and 'data consumers', as shown in the diagram. A data producer may be a business process (manual or automated), a feed from a third-party, perhaps through a web portal, or an application. On the consumer side, data is typically used by applications and other processes.



This network of suppliers and consumers is then mapped to business capabilities and the business goals that they are supporting. A cost equation on the data producer side is set against further metrics on the consumer side, which take account of frequency of data use and mode of consumption; hence, the laws of supply and demand form a 'market price'.

This is only part of the story, however.

Weightings can be applied to some data consumers (identified as 'special' consumers), to reflect their importance and encourage their use. These special consumers represent the 'elite forces' of the data world and are identified by the company to align with specific business and IT strategies. Examples of the type of data consumer that may be identified as 'special' are shown in the diagram.

The concept of special data consumers is an important one, as they add another dimension to the way data owners think about their data. Special consumers merit a hike in valuation if their owners are actively pursuing policies like security, master data and integration; but they may also be assigned negative ratings.

For example, data that has been introduced into a data repository but is not being used by any consumer is wasteful and should be eliminated. Recognising this, owners would see a drop in the value of the data for which they responsible. If data owners then address the situation, perhaps by removing the unused data through integration, their action would trigger an increase in value again.

In short, now that a tangible, monetary value has been assigned to the data repository, data owners can see how the initial data valuation is affected by their subsequent actions: initially a drop in value due to the discovery of unused data, followed by an increase if they remedy the situation - in other words, the asset valuation of data acting as a KPI.

There may be special data consumers specific to an enterprise or environment - in a financial services environment, for example, 'regulatory use'

would be a highly rated special consumer. These will need to be identified on an individual basis.

The underlying formula used to calculate data value is relatively simple; the hard work lies in establishing the data repositories that need to be owned, appointing the owners and identifying the data producers and consumers.

It is worth the effort, though.

I have identified a number of special data consumers and have given them hypothetical weightings for illustrative purposes and for use in our worked example below.

SPECIAL DATA CONSUMER TYPES

The data owner would normally encourage all potential data consumers to connect to the data repository, since according to the valuation formula this increases value. If, however, the consumer is denied data by the data owner for security compliance reasons, this should not result in a reduction of value, but should be viewed as adding a 'security consumer'. Hence, the security consumer has a hypothetical weighting of '5'.

Consumption of data by a master data implementation is the highest compliment, so the 'master data hub consumer' should be given a high weighting ('7' in my example) since it indicates that the originating data repository has one or more authoritative sources of data.

A data owner should be encouraged to reduce any duplication of data, and in doing so, the special 'integration consumer' should be invoked. The elimination of a data repository due to data and/or process integration results in the highest weighting of all - '10' in the example.

Data that is consumed into canonical messages for use in SOA or information transmission is given a positive weighting ('3' in my example) as this should encourage the use of good architectural practices and the removal of tight coupling with applications.

Data that has been introduced into a data repository but is not being used by any consumer is wasteful and should be eliminated. Data that is not used should be attributed to the 'unused data' consumer. Once it has been eliminated, it can be re-categorised as an 'integration consumer'. There may be other special data consumers specific to an enterprise or environment. These will need to be identified on an individual basis.

PART 4: A WORKED EXAMPLE OF THE CONSUMPTION-BASED VALUATION APPROACH

How could this work in practice?

In Part 3, the data producers and consumers for a hypothetical data repository were identified and given weightings to recognise certain 'special data consumers'. These included negative ratings - for unused data, in this example.

Now the formula is applied to calculate the value of the hypothetical data repository. This is:

(Initial Cost to Implement Data Producers)
The total cost of the application implementations needed to produce data. It should also contain the total cost of all data producers for the data repository whose value is being calculated.

Multiplied by
(Number of Ordinary Data Consumers)
The number of data consumers not classed as 'special consumers'.

Plus
(Number of Special Consumers, with Special Consumer Weighting applied)
The total number of special consumers, with each multiplied by their respective weightings.

A worked example:

In this example, two data producers to the repository have been identified:

1. An application which cost about £500,000 to implement in 1995
2. An external pricing feed that cost £20,000 to implement in 2005 (NB: no operating cost, inflation or depreciation has been included).

Four data consumers have been identified:

1. The data producer application is also a main consumer
2. A website that publishes the products with prices
3. A feed to the enterprise data warehouse.
4. Data access complies with current security restrictions (hence there is a security consumer) and there are no known breaches.

Applying the values to producers and weightings to special data consumers produces:

Data producer (application) = 500,000
Data producer (pricing feed) = 20,000
Data consumer (application) = 1
Data consumer (website) = 1
Data consumer (data warehouse) = 1
Security data consumer = (1 X weighting of 5) = 5

Applying the formula to calculate the value of the data repository:

(Initial cost to implement data producers)
(500,000 + 20,000)

Multiplied by
(Number of ordinary data consumers) 3

Plus
(Number of special consumers; 1

with special consumer weighting applied) 5

Data repository value is:
(500,000 + 20,000) x (3 + (1 x 5)) = £4, 160,000

Unused data is found so an 'unused data consumer' weighting of -1 is applied to each table with unused data. There are tables with columns containing unused data, showing that data is still being populated (i.e. produced) in these columns, not that the columns are empty or no longer populated.

Data repository value is:
(500,000 + 20,000) x (3 + (1 x 5) - 1) = £3, 640,000

The Data Owner takes action on the drop in value and initiates a project to stop the production of the unused data.

The Data Owner also agrees to allow customer data to be used in a customer master data project and a related master reference data project. This means that a special consumer weighting of 7 (master data hub) can be applied to each .

Data repository value is:
(500,000 + 20,000) x (3 + (1 x 5) + (2 x 7)) =
£111,440,000

Based on this simple example, the Data Owner has now put a tangible value on the data repository for which he is responsible. Moreover, he can see how this initial valuation has been affected by his subsequent actions: initially a drop in value due to the discovery of unused data, followed by an increase as he remedies the situation and then adds two further master data consumers to his list.

PART 5: CONCLUSION

Data ownership is the cornerstone of good governance; but it's not being carried out very effectively in most organisations today. I would contend that it's easier to take ownership of something that has a monetary value; something that will increase or diminish in value in response to the way it is managed.

Once business leaders can see the importance of the data, expressed in cost and consumption terms, it is more likely that 'data quality' will rise much higher up the priority stack. Rather than prioritising new infrastructure projects or upgrades to the latest desktop environment, board executives will be more motivated to 'get the data right' to support the key business strategies.

I am putting forward a practical approach to data ownership and valuation; a 'top down' approach, driven by senior and middle management. The ability to assign a meaningful value to data will not just improve data governance through more effective data ownership, but can help to focus the attention of senior management, who would perhaps otherwise dismiss data as an 'IT issue'.

Importantly, this approach also forms a first step towards the inclusion of data valuations on the balance-sheet; an area of increasing interest, particularly when valuing a company prior to acquisition.

Reply
glue

Glue Reply the Reply company which specialises in enterprise architecture, solution design and SOA to the retail market. Glue Reply provides high level, independent advice on the technology solutions that bring client's business objectives to life.

Glue Reply's core proposition is to help organisations maximise the value from their change and technology investments by helping them define, design, implements and resource best practice. Glue Reply works with blue chip and smaller companies as a trusted advisor as well as being known for getting stuck into the nuts and bolts of an it problem/project to ensure the desired outcome. Its clients include: Marks & Spencer, AXA UK, Boots, Cable & Wireless, Sainsbury's, Ministry of Defence and National Grid.

For further information: www.replyltd.co.uk - glue@replyltd.co.uk