

Data: Pure and Simple

Data is “The NEW Currency”

By

Marinus van Sandwyk

www.linkedin.com/in/mbogo

<http://commons16.sched.org/event/5aLV/21aj-data-pure-and-simple>



Your business data represents dramatic untapped commercial value. Leveraging your data in an AGILE fashion is critical for bringing new decision making capabilities to your business and new offerings RAPIDLY to the market. Is your company getting the most from its business information?

Industry thought leaders consider business data to be so important and valuable that they coined the term “The NEW Currency” to describe this most important asset.

Combine this with the disruptive technologies of “Big Data” and “Cognitive Computing” (<http://whatis.techtarget.com/definition/disruptive-technology>) and you ignore this asset to the peril of your company and career...

In this session we will discuss the value and mechanisms/methodologies to quantify the value of this crucial business asset.

We will analyse the reason applications are considered “legacy”, explain what “technical debt” is and the reason it has set in at many companies. We will also examine what impact data quality and data integrity has on executive decision-making and what realistic solutions exist to remove most, if not all, operational constraints from legacy applications. All of this can be done while leveraging the latest analytics and customer engagement models, without the usual anticipated disruption.

We will then examine an enterprise’s most valuable asset: its business information and how to leverage this in an AGILE fashion, RAPIDLY bringing new decision making capabilities to your business and new offerings to the market.



Dramatic value remains in our heritage applications, and you can modernize them at a fraction of the cost of replacement or redevelopment.

Learning Objectives:

- Discuss the asset value of heritage data
- Provide background to Big Data, Cognitive Computing and Watson Analytics and their relevance to high volume commercial OLTP
- Discuss technical debt, what caused it and how to fix it
- Explain data centric development and how legacy applications can implement a data centric architecture non-disruptively, as you perform routine maintenance
- Discuss data quality and data integrity in our legacy applications and how this prevents or complicates executive decisions
- Discuss how to implement and leverage analytics in our legacy applications
- Discuss data cleansing strategies and how to implement them
- Discuss data governance, data quality metrics and MDM (Master Data Management)

You will leave this session with a wealth of information, references, and other resources to enable you to motivate data quality, data integrity strategies and leverage your heritage applications.

Reclaim Your Heritage. Unleash Your Data Value.

Animated world spinning out of control exploding in the buzz words and terminology ...



We can all probably relate to this scenario where the world as we knew it is spinning faster and faster and exploding into new concepts and methodologies with which we need to come to grips. Clearly data is now the new currency!

These new acronyms and concepts are quite often just a “re-packaging” of old, known concepts, albeit in a different setting...

Let’s first analyse some current buzzwords and what they actually mean...

Cognitive computing – Cognitive computing is the simulation of human thought processes in a computerized model.

Cognitive computing involves self-learning systems that use data mining, pattern recognition and natural language processing to mimic the way the human brain works. The goal of cognitive computing is to create automated IT systems that are capable of solving problems without requiring human assistance.

Cognitive computing systems use machine learning algorithms. Such systems continually acquire knowledge from the data fed into them by mining data for information. The systems refine the way they look for patterns as well as the way they process data so they become capable of anticipating new problems and modelling possible solutions.

Cognitive computing is used in numerous artificial intelligence (AI) applications, including expert systems, natural language programming, neural networks, robotics and virtual reality. The term cognitive computing is closely associated with IBM’s cognitive computer system, Watson.

Data Centric Development – Data Centricity is perhaps one of the most misunderstood concepts in commercial OLTP, even whilst it has been around on the platform since the System/38, initially with application generators such as Genesis V and Synon 2E (now known as CA2E and Plex). The essence of data centricity is to make the database entirely SELF-AWARE and SELF ENFORCING.

<http://www.drdoobbs.com/architecture-and-design/data-centric-architecture-a-model-for-th/229301018>

https://en.wikipedia.org/wiki/Database-centric_architecture

Data Governance – Data governance (DG) refers to the overall management of the availability, usability, integrity, and security of the data employed in an enterprise. A sound data governance program includes a governing body or council, a defined set of procedures, and a plan to execute those procedures.

WATSON Analytics – <http://www.ibm.com/analytics/watson-analytics/us-en/>

Technical Debt – <http://martinfowler.com/bliki/TechnicalDebt.html> - I will discuss this phenomenon in quite some details a little later in this session.

Disruptive Technologies – A disruptive technology is one that displaces an established technology and shakes up the industry or a ground-breaking product that creates a completely new industry.

Here are a few examples of disruptive technologies:

- The personal computer (PC) displaced the typewriter and forever changed the way we work and communicate.
- The Windows operating system's combination of affordability and a user-friendly interface was instrumental in the rapid development of the personal computing industry in the 1990s. Personal computing disrupted the television industry, as well as a great number of other activities.
- Email transformed the way we communicating, largely displacing letter-writing and disrupting the postal and greeting card industries.
- Cell phones made it possible for people to call us anywhere and disrupted the telecom industry.
- The laptop computer and mobile computing made a mobile workforce possible and made it possible for people to connect to corporate networks and collaborate from anywhere. In many organizations, laptops replaced desktops.
- Smartphones largely replaced cell phones and PDAs and, because of the available apps, also disrupted: pocket cameras, MP3 players, calculators and GPS devices, among many other possibilities. For some mobile users, smartphones often replace laptops. Others prefer a tablet.
- Cloud computing has been a hugely disruptive technology in the business world, displacing many resources that would conventionally have been located in-house or provided as a traditionally hosted service.
- Social networking has had a major impact on the way we communicate and -- especially for personal use -- disrupting telephone, email, instant messaging and event planning.

Harvard Business School professor Clayton M. Christensen coined the term disruptive technology. In his 1997 best-selling book, "The Innovator's Dilemma," Christensen separates new technology into two categories: sustaining and disruptive. Sustaining technology relies on incremental improvements to an already established technology. Disruptive technology lacks refinement, often has performance problems because it is new, appeals to a limited audience, and may not yet have a proven practical application. (Such was the case with Alexander Graham Bell's "electrical speech machine," which we now call the telephone.)

In his book, Christensen points out that large corporations are designed to work with sustaining technologies. They excel at knowing their market, staying close to their customers, and having a mechanism in place to develop existing technology. Conversely, they have trouble capitalizing on the potential efficiencies, cost-savings, or new marketing opportunities created by low-margin disruptive technologies. Using real-world examples to illustrate his point, Christensen demonstrates how it is not unusual for a big corporation to dismiss the value of a disruptive technology because it does not reinforce current company goals, only to be blindsided as the technology matures, gains a larger audience and market share and threatens the status quo.

BIG Data – Big data is an evolving term that describes any voluminous amount of structured, semi-structured and ESPECIALLY unstructured data that has the potential to be mined for information.

Predictive Analytics – Predictive analytics is the branch of data mining concerned with the prediction of future probabilities and trends. The central element of predictive analytics is the predictor, a variable that can be measured for an individual or other entity to predict future behavior. For example, an insurance company is likely to take into account potential driving safety predictors such as age, gender, and driving record when issuing car insurance policies.

Multiple predictors are combined into a predictive model, which, when subjected to analysis, can be used to forecast future probabilities with an acceptable level of reliability. In predictive modeling, data is collected, a statistical model is formulated, predictions are made and the model is validated (or revised) as additional data becomes available. Predictive analytics are applied to many research areas, including meteorology, security, genetics, economics, and marketing.

Data Valuation – see Infonomics

Legacy – In information technology, legacy applications and data are those that have been inherited from languages, platforms, and techniques earlier than current technology. Most enterprises who use computers have legacy applications and databases that serve critical business needs. Typically, the challenge is to keep the legacy application running while converting it to newer, more efficient code that makes use of new technology and programmer skills. In the past, much programming has been written for specific manufacturers' operating systems. Currently, many companies are migrating their legacy applications to new programming languages and operating systems that follow open or standard programming interfaces. Theoretically, this will make it easier in the future to update applications without having to rewrite them entirely and will allow a company to use its applications on any manufacturer's operating system.

In information technology, legacy applications and data are those that have been inherited from languages, platforms, and techniques earlier than current technology. Most enterprises who use computers have legacy applications and databases that serve critical business needs. Typically, the challenge is to keep the legacy application running while converting it to newer, more efficient code that makes use of new technology and programmer skills. In the past, much programming has been written for specific manufacturers' operating systems. Currently, many companies are migrating their legacy applications to new programming languages and operating systems that follow open or standard programming interfaces. Theoretically, this will make it easier in the future to update applications

without having to rewrite them entirely and will allow a company to use its applications on any manufacturer's operating system.

In addition to moving to new languages, enterprises are redistributing the locations of applications and data. In general, legacy applications have to continue to run on the platforms they were developed for. Typically, new development environments account for the need to continue to support legacy applications and data. With many new tools, legacy databases can be accessed by newer programs.

Infonomics – Infonomics is a social science that involves studying the production and consumption of information and the transfer of money to produce, sell or obtain it. Simply put, infonomics is the economics of information.

Infonomics is based on the idea that information is an enterprise asset and should be counted and managed as such. Just as with economics, there are two main types of infonomics: micro and macro.

Microinfonomics focuses on information related to private individuals or businesses. Macroinfonomics focuses on the information ecosystem as a whole.

According to the information technology research and advisory firm Gartner, the principles of Infonomics have quantifiable potential which should be maximized to help budget IT and business initiatives. People who study Infonomics are called Infonomists.

Data Quality – Data quality is a perception or an assessment of data's fitness to serve its purpose in a given context.

Aspects of data quality include:

- Accuracy
- Completeness
- Update status
- Relevance
- Consistency across data sources
- Reliability
- Appropriate presentation
- Accessibility

Within an organization, acceptable data quality is crucial to operational and transactional processes and to the reliability of business analytics (BA) / business intelligence (BI) reporting. Data quality is affected by the way data is entered, stored and managed. Data quality assurance (DQA) is the process of verifying the reliability and effectiveness of data.

Maintaining data quality requires going through the data periodically and scrubbing it. Typically this involves updating it, standardizing it, and de-duplicating records to create a single view of the data, even if it is stored in multiple disparate systems. There are many vendor applications on the market to make this job easier.

SDM (Strategic Data Management) – https://en.wikipedia.org/wiki/Data_management

See especially the thought leadership articles from Bernard Marr and Jay Zaidi

Executive Decision Making – <https://hbr.org/2006/02/the-seasoned-executives-decision-making-style>

Analytics – Data analytics (DA) is the science of examining raw data with the purpose of drawing conclusions about that information. Data analytics is used in many industries to allow companies and organization to make better business decisions and in the sciences to verify or disprove existing models or theories. Data analytics is distinguished from data mining by the scope, purpose and focus of the analysis. Data miners sort through huge data sets using sophisticated software to identify undiscovered patterns and establish hidden relationships. Data analytics focuses on inference, the process of deriving a conclusion based solely on what is already known by the researcher.

The science is generally divided into exploratory data analysis (EDA), where new features in the data are discovered, and confirmatory data analysis (CDA), where existing hypotheses are proven true or false. Qualitative data analysis (QDA) is used in the social sciences to draw conclusions from non-numerical data like words, photographs or video. In information technology, the term has a special meaning in the context of IT audits, when the controls for an organization's information systems, operations and processes are examined. Data analysis is used to determine whether the systems in place effectively protect data, operate efficiently and succeed in accomplishing an organization's overall goals.

The term "analytics" has been used by many business intelligence (BI) software vendors as a buzzword to describe quite different functions. Data analytics is used to describe everything from online analytical processing (OLAP) to CRM analytics in call centers. Banks and credit cards companies, for instance, analyze withdrawal and spending patterns to prevent fraud or identity theft. E-commerce companies examine Web site traffic or navigation patterns to determine which customers are more or less likely to buy a product or service based upon prior purchases or viewing trends. Modern data analytics often use information dashboards supported by real-time data streams. So-called real-time analytics involves dynamic analysis and reporting, based on data entered into a system less than one minute before the actual time of use.

Data Integrity – Data integrity refers to the overall completeness, accuracy and consistency of data. This can be indicated by the absence of alteration between two instances or between two updates of a data record, meaning data is intact and unchanged. Data integrity is usually imposed during the database design phase through the use of standard procedures and rules. Data integrity can be maintained through the use of various error checking methods and validation procedures.

https://en.wikipedia.org/wiki/Data_integrity

MDM (Master Data Management) – Master data management (MDM) is a comprehensive method of enabling an enterprise to link all of its critical data to one file, called a master file, that provides a common point of reference. When properly done, MDM streamlines data sharing among personnel and departments. In addition, MDM can facilitate computing in multiple system architectures, platforms and applications.

The benefits of the MDM paradigm increase as the number and diversity of organizational departments, worker roles and computing applications expand. For this reason, MDM is more likely to be of value to large or complex enterprises than to small, medium-sized or simple ones. When companies merge, the implementation of MDM can minimize confusion and optimize the efficiency of the new, larger organization.

For MDM to function at its best, all personnel and departments must be taught how data is to be formatted, stored and accessed. Frequent, coordinated updates to the master data file are also essential.

“The NEW Currency” – Data currency is the monetary value assigned to data to identify its financial significance to an organization.

Once the monetary value of data assets is identified, it may be used as the unit of exchange in a transaction, either as the sole payment or in combination with money. Another reason to assign a value to data is to quantify its importance to the business. That evaluation can help an organization in data management planning and related efforts like business continuance and disaster recovery planning. This model recognizes that information assets have financial significance to the business but often involves prioritization or relative valuation rather than assigning a specific dollar figure to data.

Direct financial models of data monetization include:

- Identifying the cost of replacing data in the event of its loss.
- Identifying the amount that data contributes to an organization's revenue.
- Identifying what the data could generate if sold.

The concept of monetizing data is central to a new field of study known as infonomics, which involves the production and consumption of information and the transfer of money to produce, sell or obtain it. Infonomics is an acronym made from the first part of the word information and the last part of economics.

<http://www.businessinnovationfactory.com/summit/story/data-new-currency>

AGILITY - How organizational agility will save and destroy your company

<http://searchcio.techtarget.com/tip/How-organizational-agility-will-save-and-destroy-your-company>

Aiming for organizational agility? Prepare to change how you scope work, structure teams, develop talent, practice leadership, and deliver and sell your products.

You have probably heard of Agile methodology from your software development teams. What you might not know is that organizational agility will make the difference for your company, even beyond IT, between merely surviving and thriving.

Now, your company might survive just fine without the hard work it takes to become an agile organization. But, in my experience, the organizations that thrive in a fast-changing economy, where markets are created, destroyed and disrupted in a matter of weeks or days, are organizations that can move fast. Ready? To become nimble you will need to destroy most of what you understand about how your business works.

Why is agility so difficult to attain?

Organizational agility is talked about a lot. Indeed, in recognition that industries are changing fast, many organizations have agility in one form or another as a bullet in their corporate strategy. But if you look at most organizations, it just isn't making the huge difference that the pundits claim it will. Why is that? Business agility is hard to attain.



Before we go any further, let me define "agile organization."

If you are able to adapt to the changes in your business context (customers, technology, markets, and so on) as fast as, or faster than, they occur, you are an agile organization.

That means being able to answer the following questions:

- How fast are changes happening in your industry?
- How quickly are your competitors delivering new products and services or new features and functions?
- How quickly are your customers' needs changing?
- How quickly does your underlying technology change?

These elements are compounding. Meaning, if you have a change in market segmentation or new market creation and your underlying technology changes, your organization has to adapt to both changes simultaneously, compounding the complexity of the change and making it all the more important that your organization is nimble.

With this definition of organizational agility in mind, let's take a look at your business systems, focusing in this column strictly on product or service creation. (Maintenance and support of products and services are also impacted, but covering all that would make this a long article. We also will not cover finance and budgeting here, but they are impacted too.)

Here are six key areas of product creation that are impacted by the need to be nimble:

How you scope work

Most organizations conceive of work in large batches that are worked on for a long time and then delivered to their customers. But this is a huge risk when things change so quickly. What happens if you are in the middle of the creation of a product and your competitor comes out with an innovation that changes your market? What happens if your underlying technology changes mid-development and you are left developing on an outdated platform?

The contrary model is to conceive of your work in small increments with a direction in mind but with fast delivery of small, autonomous elements of your product or service, that provide value to your customers, but that may not be as "complete" as you normally would conceive of a product or service. Most often, your customers don't want or need all the complexity you build into your offerings. Create an offering that has the minimum necessary to provide the highest value to your customers and then gauge their feedback. This is a fundamental difference in how you think about work. It impacts how projects/products are created, it impacts governance, it impacts finance, it impacts team structure and it impacts delivery and marketing.

How you structure your teams

You probably have heard of the Tuckman model of team development. That is the one with the four stages: Forming, Storming, Norming and Performing. Teams, in one way or another, go through this process. To a small degree, they go through it whenever a team member is added or removed. When at least 30% of the team has changed, they definitely go through the cycle again. This process slows the delivery of work because people are focused on the social/team dynamic rather than the work. This is a delay that we would like to avoid.



There is another option. Many people phrase the contrast this way: The traditional approach is that we have work and we bring "the right people" to the work. To be nimble -- to have organizational agility -- we have to have great teams of people and bring the work to the team.

How you develop people

To support this approach of having great, long-lived teams and bringing the work to them, we need a slightly different approach to developing our people. Who you want is known as a T-shaped person. Yes, you want people who are experts in their area, but they also need to have broad expertise. They need to be able to pick up any work that needs to get done and help move the work along.

One client had a bench one deep and that one person decided to retire. The next six months were a scramble to try to transfer 40 years of experience to a few people to support the system after that employee left. Don't be that organization.

Traditional models for staff development involve helping them deepen their expertise in their area of specialization. Think instead like a nimble startup. What do you want if you can only hire a few people? Do you want experts in small areas? No. You want people who can pick up the ball and run with it. You want to develop these T-shaped people. Cross-training in a number of areas is just as important as development of specialization.

How you lead and manage people

If your organization is going to be nimble, then leadership decisions need to be made as quickly as possible so that real action can be taken. This is a shift from a focus on making all the right decisions and reducing failure to making small, quick decisions, learning and adapting to quick failure. This scares so many executives. We have been trained since childhood to reduce failure. Our papers in school were graded once, and if you got it wrong, that was it; you had no second chance. It was really a focus on timing rather than learning. This was reinforced throughout school and in our jobs after college. Today in school, there is a greater focus on learning than on timing. My wife is a middle school teacher, and in her class, students can turn in a paper as many times as they like, making the grades as excellent as the students choose. She wants to give them quick feedback and then help them improve.

This same model applies to how you lead and manage people. You need to give them guidance and direction but then set them free to make quick decisions without asking you for approval on every decision.

Your delivery pipelines

To achieve organizational agility, your product and service delivery pipeline must be reconstructed to handle the increased cadence of change. So, if your change cycle in your industry and for clients is six months, but your delivery pipeline is 12 months, you have some work to do. What people don't recognize when they are faced with this kind of shift in timeline is that it comes with a commensurate change in scope. You will be delivering half as much product or service change. The impact on your customer is half. The impact on your support staff is half. The amount of customer documentation is half.

The simplest thing to do is to look at your current offering line-up and cut it in half. Don't do this arbitrarily; rather, attempt to deliver the most important things to your customers first. Put the less important things in the second delivery. This way you get the biggest bang for your buck.

Sales and marketing

If you are delivering faster, your marketing and sales teams also need to be working twice as fast. They need to be announcing new products and services twice as fast. They need to be updating their sales knowledge twice as fast. Marketing campaigns need to be modified. But again, remember that the new offering is only going to be half the scope that it was before, so marketing personnel doesn't need to communicate as many new features or options because there is only half as much change.

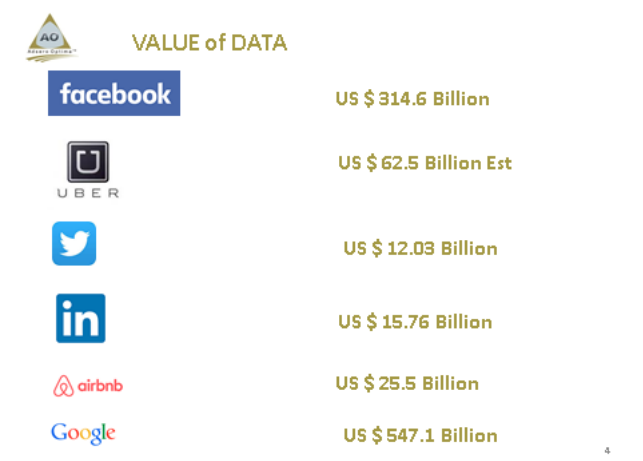
I have worked with a couple of marketing teams, and they were able to make this change, and it was just as complex as it was for the product development teams. Don't underestimate the impact on sales and marketing.

Becoming nimble -- becoming an agile organization -- isn't just an IT or software concern. Every organization today needs to assess if they are nimble enough to lead their market. Because organizational agility fundamentally changes how you conceive of your work, it impacts the whole organization, not just a delivery of a particular product. If your organization is not nimble enough to keep up with the market and your competitors, there is something you can do about it! It won't be easy, but it will be worth it.

About the author:

Joseph Flahiff is an internationally recognized leadership and organizational agility expert at Whitewater Projects Inc. He has worked with Fortune 50 and Fortune 500 companies, government agencies, start-ups and publicly traded firms, where he has been lauded by executives as an experienced, pragmatic and innovative adviser. He is the author of *Being Agile in a Waterfall World: A practical guide for complex organizations*. Learn more at www.whitewaterprojects.com; to contact Joseph send an email to joseph@whitewaterprojects.com

<https://www.linkedin.com/in/josephflahiff>



This slide should change your perception of the real VALUE of DATA conclusively and for the remainder of your career...

The BULK of the market value of these organisations are dictated by DATA, and not “old” or established principles of valuation (EPS, NAV, NPBT, EM, etc.)... The remarkable spread and differences in the various valuations we shared here, highlights how much the principle of “Infonomics” and valuation of DATA as an asset is evolving... It will remain a hotly debated subject for a long time to come, as the fundamental principles are debated and mature...



INFONOMICS – Data as an ASSET

Non-financial methods

1. **Intrinsic value of information.**
2. **Business value of information**
3. **Performance value of information.**

Financial methods

1. **Cost value of information.**
2. **Economic value of information.**
3. **Market value of information.**



Doug Laney – Gartner
<https://www.linkedin.com/in/douglaney>

3

The thought leader and company to follow on this evolving and maturing concept are Doug Laney and his partners at Gartner. The work they have done in recent years is significant.

Please see some of my links imbedded in this document and provided on the AO Website. This is a MASSIVE subject, which can easily consume days of debate and which is way beyond what I am trying to share fundamentally with you today.

In a nutshell, the following approaches MAY PERHAPS provide you with some mechanisms to perform some of your own valuations...

Non-financial methods

1. Intrinsic value of information. This model doesn't "take into account the business value at all," Laney said, but focuses instead on the data's intrinsic value. The model quantifies data quality by breaking it into characteristics such as accuracy, accessibility and completeness. Each characteristic is rated and then tallied for a final score. Laney, who teamed up with Gartner's Ted Friedman to quantify a dozen data quality characteristics, includes scarcity in the equation. "Data that's more unique to your organization and not available to your competitors or the larger marketplace, we believe, has the potential to provide more value to you," Laney said. As with any of the six models, this one can be tailored to the company, which could, for example, "assign weighting factors" to each characteristic, he said.

2. Business value of information. This model measures data characteristics in relation to one or more business processes. Accuracy and completeness, for example, are evaluated, as is timeliness "because even if data is relevant to a business process, if it's not timely, how valuable is it really?" Laney said. The model can be tailored to fit the organization's needs and even applied to specific data types such as unstructured data or third-party data.

3. Performance value of information. This model is "much more empirical in nature" because it measures the data's impact on one or more key performance indicators (KPIs) over time, Laney said. Take the sales department, for example. "If your salespeople had access to competitor pricing data, how much quicker could they close sales?" Laney said. Businesses can run an experiment by comparing how a control group with no access to competitor pricing data performs against an experimental group. Or, if businesses have

neither the time nor the ability to run an experiment, they can substitute proxy data for control group data, he said.

Financial methods

4. Cost value of information. This model measures the cost of "acquiring or replacing lost information." After the Sept. 11, 2001 terrorist attacks, as clients began calling Laney to figure out how to recoup from, in some cases, a total loss of data, they developed a method to quantify information's value based on what accountants "refer to as 'replacement costs,'" Laney said. A value is assigned to the data by measuring lost revenue and how much it would cost to acquire the data. "This is the way valuation experts value most intangible assets that don't have a discernible market value or are generating a market stream," Laney said.

5. Economic value of information. This model measures how an information asset contributes to the revenue of an organization. "This is our KPI model again, but instead of any given KPI, we're looking at revenue," Laney said. To illustrate his point, he returned to his sales example. An experimental group is given access to competitor pricing data and a control group isn't. "Instead of looking at time-to-sale, we're looking at revenue generated by any given salesperson" over a given period of time, Laney said. "That will give us a good sense of the value of that data." CIOs should factor in the cost it takes to acquire, administer and "bake that data into the system the salespeople are using," he said. They should also consider the data's life span. Competitor pricing data, for example, has a shelf life, which should be factored into its value.

6. Market value of information. This model measures revenue generated by "selling, renting or bartering" corporate data, which Laney considers to be one of the best ways to value an information asset. The problem is, most information assets don't have what accountants call an "open arms-length market," or what the price of the data would be on the open market, according to Laney. A way around this is to figure out what similar data from syndicated data providers or competitors is going for. After determining the data's premium price, Laney suggests figuring out what he calls a "discount value." "When we sell data, we're not really selling it," he said. "We're licensing it." The discount rate will vary based on the number of times a company sells the information and other factors. "But, again, it's not the value that's important," Laney said. "It's tracking over time."

Principles of infonomics

1. Information is an actual asset (if not a recognized asset class).
2. Information has both potential and realized value.
3. Information's value can be quantified.
4. Information should be accounted for as an asset (internally).
5. Information's realized value should be maximized.
6. Information's value should be used to help budget IT and business initiatives.
7. Information should be managed as an asset.

Source: Gartner Inc.

"Imagine the difference if I tell you that you're in charge of our customer database versus you're in charge of our \$50 million customer information asset," Laney said. "It's a bit of an attitudinal change that can take shape."

Right now, data is siloed, even duplicated in some cases, and multiple versions of the truth are common. "How do we look across the entire enterprise and get everyone on the same page with all of the data?" he

said. Embracing Infonomics, he believes, may help by making an intangible concept -- the value of data to the organization -- tangible -- Jimm Johnson, Scripps Health

<http://searchcio.techtarget.com/feature/Infonomics-treats-data-as-a-business-asset>

<http://searchcio.techtarget.com/feature/Six-ways-to-measure-the-value-of-your-information-assets>



6

Now that we have LIKELY (depending on your own assessment, considering the facts around your applications) established that DRAMATIC value is captured within your heritage applications, you need to ask a CRITICAL question about the QUALITY and INTEGRITY of your data, especially if you system/application has been used for decades.

How many orphan records do you have in your database? How much duplication and inconsistencies exist, especially due to our use of the old programming model (OPM) in the 1980's and early 1990's, prior to the maturity of DB2 for i (OS/400 at that stage) and the introduction of the ILE programming model... This subject is closely linked with the issue of "Technical Debt", which I will discuss in some detail a little later on...

This is an extremely important consideration, as data links every member off your staff; it's the glue that holds the business together...

Growing the business means increasing profit, and it should mean improving efficiency. The proper management of data should be central to your plans. If you expand without handling data properly, you'll end up in a mess, where no single department fully understands how their management of data affects everybody else.

Data quality isn't a natural occurrence, and achieving a high standard of accuracy doesn't occur overnight. If you want to ensure that your data is pure, expect to invest in it.

Focus on Quality Over Quantity

Focus on collecting high quality data, and there are two positive outcomes; your ROI increases, and the cost of storage and management is reduced. Big data sounds great, but small (and structured as opposed to unstructured) is fine for SMEs.



In fact, the results can be FAR more disastrous...
 ...it may cost you your job... or worse...
 The demise of your company...

Data in a database rots and decays in exactly the same way as teeth do. Like frequent brushing of teeth, frequent data maintenance is required to keep the data in good health, ensuring that the rot cannot progress to a catastrophic stage. That's one good argument for data maintenance, and it proves why it is an unavoidable task that all businesses must commit to.

When we think about data, we can compare it to caring for our health. In particular, data maintenance is a lot like brushing your teeth. We brush our teeth at least twice a day to stop decay from taking hold. If we didn't, the sugar that we consume would gnaw away at the enamel and cause rot to set in. The longer we leave it between brushings, the more vulnerable our teeth become. Similarly, our database must be continually cared for and maintained.

MUCH of this CRITICAL process can be left to the database engine, provided that you have adopted GENUINE data centricity implemented at the database layer and TIGHT data validation rules enforced within the database layer...



...data quality problems cost U.S. businesses more than

\$600 billion *pa* 2002

... **current impact** of bad data and data quality issues on US businesses estimated at

\$1 Trillion ...or

6% of US GDP!



Impact of bad data and data quality issues

In 2002, The Data Warehousing Institute (TDWI) published a report titled "Data Quality and the Bottom Line: Achieving Business Success through a Commitment to High Quality Data.". *This report showed that there is a significant gap between the perception and reality regarding the quality of data in many organisations, and that data quality problems cost U.S. businesses more than \$600 billion a year. The report's findings were based on interviews with industry experts, leading-edge customers, and survey data from 647 respondents and primarily focused on the impacts of bad customer and address data.* Imagine

what the current costs are, 13 years later, since data volumes have exploded due to social and mobile applications and data is far more complex, with respect to variety.

My highly conservative projection puts the current impact of bad data and data quality issues on US businesses at One Trillion Dollars. To put this in perspective, the US Gross Domestic Product (GDP) is approximately 17 Trillion Dollars. So, at a minimum, bad data is costing the U.S. 6% of its GDP. The worldwide impact may be in the Trillions. These metrics should get the attention of Policy Makers, Shareholders and Boards of Directors!! - Jay Zaidi: "Bad Data is Costing the U.S. At Least 6% of its GDP" by - <https://www.linkedin.com/pulse/bad-data-impacting-your-organisations-bottom-line-jay-zaidi>



??? Metadata???

What is Metadata?

9

Metadata is probably the most important concept in data management. But, it is also the most misunderstood concept and therefore not given the importance it deserves.

Metadata is a rather confusing and technical term, so I (Jay) prefer to call it "Contextual Data". I believe "Contextual Data" provides a much simpler and clearer meaning and is easier to understand. So, going forward I will use these terms interchangeably.

The essence of Metadata is that it is information (attributes) describing your data, such as length of a field (column), the nature of that component (zoned, packed, numeric, alpha numeric), other information such as validation rules, behaviour, URL's, text descriptions, associated icons, associated JSON, etc., etc. in short any information describing how that data should be presented, manipulated, controlled, etc.

Business Glossaries and Metadata – Governance of Enterprise Reference Data – Why Do We Care?
By Lowell Fryman – <https://www.linkedin.com/in/lowell-fryman-cbip-cdmp-152a613>

<http://tdan.com/business-glossaries-and-metadata-governance-of-enterprise-reference-data-why-do-we-care/19525?>

Reference Data Management often struggles to be governed as an enterprise data asset. Many organizations still view Reference Data as an individual application problem and thus Reference Data is often an application afterthought. The Data Governance team has to ask questions such as:

- Who is accountable to source, maintain, and distribute reference data?
- Are the domain data values isolated to one application, or is it common across multiple applications and business units?

- What are the implications for Data Consumer usage, management reporting, analytics/Big Data, and regulatory reporting?
- Is the management responsibility at a business unit, Finance, an application, the CIO or more common now the CDO?

For most organizations the answer is “it depends.” That is not a good answer to tell your audit team or industry regulators. Today Reference Data should be considered as a critical enterprise data asset. Not just for an individual business unit or application, but as an enterprise asset and governed as such.

Clarity is Important

Like all critical data assets, having a solid description of the language and terminology is important. First, what do I mean by the term Reference Data? Reference Data is data that defines or creates a context for other data. For example, Reference Data helps to identify characteristics of a Customer. Such characteristics as what industry the Customer is in, what country the Customer’s account was established in, where the Customer lives, where we send the Customers bills to, and the gender of the Customer.

I can remember when we did not have the concept of Reference Data. We created the concept so we no longer had to change the application code anytime a new valid value was needed to manage business processes, or even when the business rules for data content validation changed. Reference Data is used in the business rules for the capture and management of data content; in the “dropdown boxes,” in the data quality rules, validation of the data contents, in data integration aggregations (data warehouse), as well as “roll-ups” and aggregations for Analytics and reporting including Big Data analytics.

The underlying issue may be that Reference Data is not always considered to be business data. Many IT and business operations individuals still view Reference Data as application level metadata. That is not, and never has been the case.

Some examples of Reference Data used to manage business rules and data validations include:

Places – Location Data (Geographic and Political data)

Continent, Region, Country, State/Province, County, City, & Postal code
Census Tract, Economic Designation, Political Designation, Zoning Codes, ...

Things

Monetary – Currency Types & Codes, Exchange Codes, Transfer Codes, & Account Type
System or enterprise specific objects – GL Accounts, Cost Codes, Industry code, & Organization Hierarchy

People

Customer Type, Contact Type, Account Status Code, Risk Rating, Occupation Type
Counter-Party, Supply Partner, Industry Code, Risk Ratings, Relationship Codes

Challenges of Reference Data

The governance of Reference Data has a number of challenges in most organizations. Let’s see which of these you can relate with.

- Lack of accountability or ownership as a data asset. “Just make up something for the application to work.”
- Confusion and/or conflict over accountability and sourcing responsibilities. “My business unit and budget should not be responsible for managing and supplying this data to the rest of the enterprise; you need it, you go get it.”

- Perception that Reference Data is only an IT problem to resolve. “Just add the value to the dropdown box.”
- Reference Data never changes so we don’t need a management infrastructure and processes. “We just need to get through QA testing and it never changes anyway.”
- We all know what it means so why define it. “It’s just a loan application status code.”
- No Data Catalogue or Sharing Agreements for Data Consumers to know what is available and how to use it. “Go look at the dropdown box and the application user manual done in 1994.”
- Reference Data has no relationships; it is flat with no dependencies or hierarchies. “All the metadata values are independently entered by the coders.”
- Very little (let’s be honest it’s none) data quality testing is done on Reference Data prior to its implementation in production. “It’s just a select box value.”
- There is a lack of Data Governance processes over the life-cycle of Reference Data. “What life-cycle; it’s not business data.”
- Reference Data values used across different business units or in Analytics and reporting can be different than those used in operational applications. “My hierarchies are better and more accurate than yours; but mine are for G/L accounting.”

Wonderful, So What Steps Do We Take to Govern Reference Data!

As always, we can’t start the “12 step program” until we accept that we have an issue and we want to resolve that issue. Both recognition and desire to fix the issue are critical. Nothing can actually happen until we have the desire to fix the issue. I suggest that it is the responsibility of the CDO or Data Governance team to raise the issue and drive recognition that a solution and infrastructure must be established. Acceptance of the issue and the drive to fix the issues is often considered an organizational change or cultural change challenge. Those concepts and actions are not new to most organizations so leverage what has worked in the past. The solutions for Data Governance issues like this are generally driven by a cross-functional Data Governance Committee. Awareness and communication across the organization are critical to resolve the challenges. The following is my 12 step program.

I suggest that the Data Governance principles state that “all Reference Data will comply with the policies and operating practice of Data Governance (yes let’s start with the desire that “all” is an enterprise asset). Generally, we see the principles simply stated such as “all Reference Data will be managed as an enterprise asset similar to the management of our data content.” The policy level can then allow for some Reference Data to be managed in a centralized infrastructure while certain Reference Data is managed as a federated infrastructure. The Reference Data Policy can then define and describe the expectations and standards for Reference Data.

Accountability and Stewardship roles and responsibilities should be identified and accepted. This will be by the type of Reference Data. For example, your Commercial business unit may be determined to be accountable, stewards, and the authoritative source for NAICS codes. They will assume the roles and responsibilities associated with NAICS codes.

Harmonized business terms, definitions and understanding for Reference Data will be across the enterprise and managed in the Business Glossary. Each type of Reference Data should be identified in your Business Glossary along with the valid value set, and business rules. These will be for example, Country Code, Currency Code, and Standard Industry Classification Code. These processes for identifying, harmonizing, and management of business terms are complex, but I’m keeping it to one step in this list.



Periodic data profiling is used to validate the Business Term and Business Definition of Reference Data. Data profiling should be used to validate the valid value set identified for the Business Term. The Business Term Name and Definition should be applicable for the value set implemented.

Each type of Reference Data will have life-cycle management processes and infrastructure roles and responsibilities defined. The Reference Data acquisition processes, maintenance, timing and architecture must be well defined and communicated. One and only one source should exist for each type of Reference Data. It is fine if your architecture is centralized for some Reference Data and federated for other data.

Authorized Sources and Authorized Distribution processes and systems must be identified (preferably in a Data Catalog). People need to know the definition, availability, where to source, and how to source the data they need.

Data Quality management is used to ensure the defined valid values and business rules in the Business Glossary are physically implemented in the Reference Data Management infrastructure. A Data Quality dashboard is important for the Data Consumer community. Remember this dashboard will be limited in scope to the Reference Data, not the data content occurrences using the Reference Data.

Data Consumers are identified and included in the accountability roles and responsibilities. This is a step that is an on-going process. The number of Data Consumers should expand as your Reference Data usage grows. Teams will come and go, but the accountability for the usage of specific Reference Data should not change. The accountability of Data Consumers usage of data is often over looked.

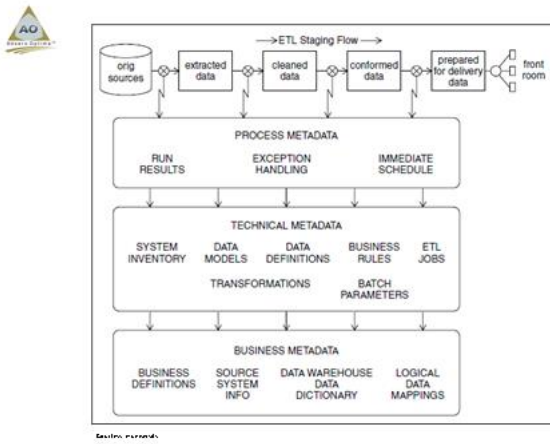
Data Sharing Agreements should be established for consistent Reference Data usage including the processes for aggregations and “roll-up” hierarchies. Hierarchies must be considered a component of the Reference Data, as applicable. Two different valid value sets for hierarchies for the same Reference Data must be considered as two different concepts and managed separately. Else we have semantic confusion and reporting differences. Different Business Terms for the different hierarchies are likely important to identify.

Issue management and issue escalation processes are needed for Reference Data Change Management. However, Reference Data does not need an issue management process different than the rest of your data. Use the processes that you have for the issue management for all data.

Reference Data must be available to all authorized Data Consumers, either individuals or applications. Multiple distribution formats are likely necessary (such as Web service, XML or flat file distribution). While Reference Data may not have specific security or data protection requirements, you will likely have Data Retention standards applicable for each source. The Data Governance team can help identify and resolve these requirements.

Reference Data management is far more important and complex in most organizations as we consider the Big Data analytics, management/audit reporting and regulatory reporting environments we are operating under today. The Reference Data used in our operational applications has to be integrated with the data aggregation and hierarchies used in our reporting applications. Most of those are very separate applications today. Great Reference Data Management and Governance is critical to enable the alignment, as well as reduce the cost of business operations. Your Business Glossary will be an effective enabler of your Reference Data program. It is OK, stay calm and allow your Business Glossary to prosper.

https://en.wikipedia.org/wiki/Meta-data_management



Let me elaborate a little and provide an example, to make this even clearer. Different types of "Contextual Data" such as Business Metadata, Operational Metadata, Process Metadata, and Technical Metadata can be captured about specific data elements or events. The diagram (courtesy [e-university](#)) provides a graphical representation of the different types of metadata that can be captured, as data is routed via an extract, transform and load process (although we believe with the power of our machines these days, that ETL processes are no longer required and that you should analyse your LIVE data directly, which means your analysis data can no longer "age").

Business Metadata is typically captured in Data Dictionaries and Glossaries, Technical Metadata is captured within data modelling tools, database management systems and Extract, Transform and Load (ETL) tools and Process Metadata is captured in various other systems. All this contextual information is very beneficial to Programmers, Database Developers, Analysts and Data Scientists in their daily tasks.

However, please note that the management of metadata and master (or strategic) data management is probably one of the fastest evolving technologies and subject matter, due to the pressure "Big Data" and extracting maximum ROI from our data resources.

Some of the other uses of Contextual Data are:

Contextual is critical for processing Big Data, since Hadoop is based on a "Schema on Read" paradigm and therefore has no contextual information about the data on read,

Contextual data is required to better define data quality requirements,

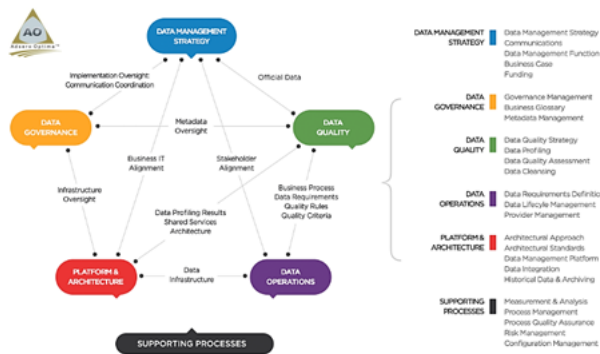
Contextual data helps in data discovery and search functions - something that most Business and Data Analysts and Data Scientists do as part of their data discovery and analysis work.

The purpose of all this metadata is to drive discoverability and increase sales,

Contextual data is used to annotate reports and in analytics, and

Contextual data can also inform data consumers of the data about who is accountable for it from a governance perspective, which business processes use it and which data domain it belongs to.

Contextual Data should be a critical component of every organisation's data management program. Establishing a Metadata Repository, seeding it with contextual data and maintaining the data over time, will provide significant benefits.



Courtesy: Jay Zaidi

Strategic & Master Data Management (MDM)

11

Tools are enablers and accelerators, but will not solve business problems - if there is poor management, teams don't leverage the tool, there is a lack of skilled data management staff, the tools aren't integrated and aren't customised for use.

A Metadata or Context Management tool may require significant upfront effort to configure, upload content and keep the content current. Given the massive amount of contextual data that organisations store, there is ongoing investment and governance required.

The bottom line is that tools enable organisations to automate many of the manual processes. They can also provide valuable insights, enable data discovery and perform other functions, which results in a significant return-on-investment.

Data quality is CRUCIAL and here's why -

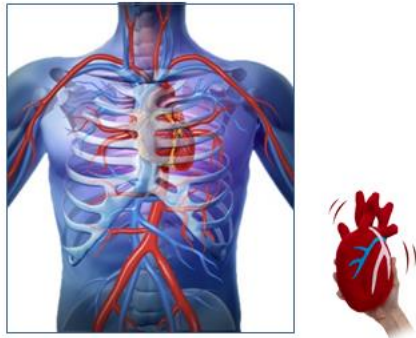
Decisions: Organisations rely on data to make critical business decisions. Using low quality data results in low quality decisions - not something that is desirable,

Policies: Business leaders develop policies based on data. Basing policies on low quality data results in poor policies that aren't optimal, and

Reports, Forecasts and Projections: Data is the raw material used for financial and operational reporting, forecasting and projections that feed into strategic planning.

Data quality matters and organisations need to focus on it, if they want to make better decisions, develop effective policies and want to improve their strategic planning processes.

Reference: Jay Zaidi



Business data - lifeblood of your organisation!

52

It is CRITICAL to acknowledge that your business data is the lifeblood of your organisation and that YOU hold that in the palm of your hand...

I would offer that the primary reason businesses aren't able to achieve their full potential is because business leaders aren't "Standing up for Data". I don't want business leaders to 'Lean In', but to 'Stand Up' for data.

What does "Standing up for data" means?

It means that business leaders must do the following 5 things:

1. *Treat data as a strategic asset.* Give it more importance than physical assets such as buildings and infrastructure,
2. *Take ownership for the governance of their data assets,* rather than delegating this important function to the Information Technology organisation or some other team,
3. *Invest in the quality of business critical data,*
4. *Clearly articulate the questions that they want answered* - and challenge their IT and Data organisations to use all the data assets to answer them, and
5. *Always ask these questions when they are making decisions based on data:*
 - Where did this data come from?
 - Who is accountable for the definition and governance of this data?
 - What's been done to this data (i.e. transformation)?
 - Does this data meet our quality requirements and what are the data quality metrics associated with it?



Impact of unstructured data

13

It is important to recognise what impact Big Data (aka unstructured data) will have on your management burden and the investment requirements to retrieve valuable information/insight from this. It is significantly less expensive to retrieve insight from structured data...

20+ Mind-Blowing Facts About Big Data Everyone Must Read – Bernard Marr

<https://www.linkedin.com/pulse/20-mind-blowing-facts-big-data-everyone-must-read-bernard-marr>

Big data is not a fad. We are just at the beginning of a revolution that will touch every business and every life on this planet.

But loads of people are still treating the concept of big data as something they can choose to ignore — when actually, they're about to be run over by the steamroller that is big data.

Don't believe me (Bernard Marr)? Here are 20 stats that should convince anyone that big data needs their attention:

1. The data volumes are exploding, more data has been created in the past two years than in the entire previous history of the human race.
2. Data is growing faster than ever before and by the year 2020, about 7 megabytes of new information will be created every second for every human being on the planet.
3. By then, our accumulated digital universe of data will grow from 4.4 zettabytes today to around 44 zettabytes.
4. Every second we create new data. For example, we perform 40,000 search queries every second (on Google alone), which makes it 3.5 searches per day and 1.2 trillion searches per year.
5. In Aug 2015, over 1 billion people used Facebook in a single day.
6. Facebook users send on average 25 million messages and view 2.77 million videos every minute.
7. We are seeing a massive growth in video and photo data, where every minute up to 300 hours of video are uploaded to YouTube alone.
8. In 2015, a staggering 1 trillion photos will be taken and billions of them will be shared online. By 2017, nearly 80% of photos will be taken on smart phones.
9. This year, over 4 billion smart phones will be shipped - all packed with sensors capable of collecting all kinds of data, not to mention the data the users create themselves.

10. By 2020, we will have over 1 billion smartphone users globally (overtaking basic fixed phone subscriptions).
11. Within five years there will be over 50 billion smart connected devices in the world, all developed to collect, analyze and share data.
12. By 2020, at least a third of all data will pass through the cloud (a network of servers connected over the Internet).
13. Distributed computing (performing computing tasks using a network of computers in the cloud) is very real. Google uses it every day to involve about 1,000 computers in answering a single search query, which takes no more than 0.2 seconds to complete.
14. The Hadoop (open source software for distributed computing) market is forecast to grow at a compound annual growth rate 58% surpassing \$1 billion by 2020.
15. Estimates suggest that by better integrating big data, healthcare could save as much as \$300 billion a year — that's equal to reducing costs by \$1000 a year for every man, woman, and child.
16. The White House has already invested more than \$200 million in big data projects.
17. For a typical Fortune 1000 company, just a 10% increase in data accessibility will result in more than \$65 million additional net income.
18. Retailers who leverage the full power of big data could increase their operating margins by as much as 60%.
19. 73% of organizations have already invested or plan to invest in big data by 2016
20. And one of my favourite stats: At the moment less than 5% of all data is ever analysed and used, just imagine the potential here.

My (Bernard Marr) prediction? At the rate at which data and our ability to analyze it are growing, businesses of all sizes — large and small — will be using some form of data analytics to impact their business in the next five years.

The question isn't whether or not big data is here to stay; **the question is are you ready??**



Data - a strategic asset

14

Business leaders have a vested interest in ensuring that data is treated as a strategic asset, that business teams take accountability for it and that there is ongoing investment in data management and analytics.

Why specifically focus on 'Business Leaders'? *It's because business leaders like the CEOs, CMOs, CDOs, CFOs are on the front lines* - developing business strategies, driving product innovation, and selling products and services - constantly working to gain competitive advantage and drive shareholder value. This is no mean task and they need every arrow in their quiver - data being one. What they have to realise is that "*Data is their friend*". They shouldn't shy away from it, but embrace it with open arms, invest in it, take ownership of it and be accountable for its governance.



Mismanagement of data resources!

15

If you do not pro-actively PLAN to manage your structured and unstructured data resources, we have little doubt that you WILL drown...

For SME's specifically, FOCUS on your STRUCTURED data, and then SELECTIVELY, gradually and cautiously AUGMENT your analysis and insight, with Big Data, exploring Watson... More about this a little later...

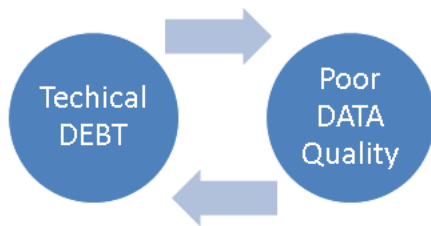


STRUCTURED ~~Versus~~ AND UNSTRUCTURED DATA

Unstructured data augments structured data

16

It is especially CRUCIAL for you to recognise the IMMENSE value in your structured data and that you can retrieve and access trends in a significantly easier manner from structured data, than from unstructured data. Also, that unstructured data DOES NOT compete with structured data, but AUGMENTS it.



Technical Debt & Poor Data Quality closely related

17

It is extremely important to recognise that there is an exceptional close relationship between “Technical Debt” and “Poor Data Quality”. As indicated, I will discuss technical debt shortly, to highlight why this phenomenon compromises data quality. This can be corrected exceptionally efficiently, completely non-disruptively and at a fraction of the cost of replacement...



Andrew White – Gartner



Howard Baldwin – Independent Analyst



Bernard Marr – API



Jay Zaidi – AlyData



Doug Laney – Gartner



Martin Doyle – DQ Global

Infonomics & data quality thought leaders

18

Quite a bit of the material that I have used so far in this discussion, originated in some form from other thought leaders in the industry that specialises on this extremely important subject. Therefore, I would suggest if you decide to research this subject a little more, to start with the following “Infonomics and Data Quality Industry Leaders”.

Andrew White – Gartner

<https://www.linkedin.com/in/andrew-white-0516a>

Alex Bakker – Saugatuck

<https://www.linkedin.com/in/alex-bakker-986a4018>

Doug Laney – Gartner

<https://www.linkedin.com/in/douglaney>

Howard Baldwin – Independent Analyst & Author

<https://www.linkedin.com/in/howardbaldwin>



Jimm Johnson – Scripps Healthcare

<https://www.linkedin.com/in/johnsonjimm>

Martin Doyle – DQ Global

<https://www.linkedin.com/in/martindoyle>

Jay Zaidi – AlyData

<https://www.linkedin.com/in/javedzaidi>

Bernard Marr – Advanced Performance Institute

<https://www.linkedin.com/in/bernardmarr>



Don't trash your heritage!

19

Most of you are probably asking by now “What is the relevance of all this talk about DATA, the value of data and the quality of data?...

Fact is, you are sitting on the MOTHERLODE of gold... your heritage application assets and ESPECIALLY your DATA...

Most companies are grappling with this difficult issue: modernize or replace.

Many consider dumping their legacy, their heritage, for the latest “silver bullet” solution, which is tantamount to dumping a 1963 Chevy Corvette in mint condition, on the scrap heap simply because it is considered old!

The purpose today is to inter alia discuss this and provide guidelines to help differentiate between fact and fiction regarding legacy (we prefer heritage) applications.

And to show how to preserve and add new life to your heritage applications which will provide a solid foundation for so much more.

The fundamental question you and your organization have to ask yourselves, is how much value your legacy system delivers to your company and your customers? Additionally, you need to ask what the cumulative value of the intellectual capital invested in your applications is (this is usually much more than what you think initially – quantify it for yourself). Most important though is to ascertain what competitive advantage your current system provides.

Often business will respond that there is little to no value in their legacy application... If that is true, why then is the company successful??

In every single instance we have been involved with internationally, the business leaders quickly admitted that the functionality delivered by their legacy LOB (Line-of-Business) application was of significant value, but this was negated by severe operational constraints, most significant of these; AGILITY.

In our experience, there remains MASSIVE value in your “legacy” applications and this can be leveraged for dramatic results and benefit to your company, relatively easily. It is entirely feasible to extend the life of these applications for another 10 to 20 years, whilst remaining competitive!!

The use of the word “legacy” in IT as implying something “bad” is truly something I grapple with. Legacy in ANY other context is something valuable and good. In our opinion, it is time that we as IT stand up and challenge this notion. There is ENORMOUS value in our legacy, provided that we remove the constraints or shackles. As a result, we prefer referring to our heritage and highlighting to business how this heritage has made us successful. Look at the formal definition:

legacy

1. (Law) a gift by will, esp of money or personal property
2. (Law) something handed down or received from an ancestor or predecessor

leg·a·cy

1. *Law* . a gift of property, especially personal property, as money, by will; a bequest.
2. anything handed down from the past, as from an ancestor or predecessor: *the legacy of ancient Rome*.
3. an applicant to or student at a school that was attended by his or her parent.

her·i·tage

1. Property that is or can be inherited; an inheritance.
2. Something that is passed down from preceding generations; a tradition.
3. The status acquired by a person through birth; a birthright: *a heritage of affluence and social position*.

her·it·age

1. **something** that comes or belongs to one by reason of birth; an inherited lot or portion: *a heritage of poverty and suffering; a national heritage of honor, pride, and courage*.
2. something reserved for one: *the heritage of the righteous*.
3. *Law*.
 - a. something that has been or may be inherited by legal descent or succession.
 - b. any property, especially land, that devolves by right of inheritance.



IBM i a powerful vehicle ...

20

The fact of the matter is that we truly have the best commercial online transaction processing (OLTP) platform in the history of computing, yet our users quite often perceives this as...:



Why, then, is your system perceived as “legacy”?

21

The old adage of “perception is reality” is accurate here. Regardless of what we know or believe, our users perceive our “top of the line” sports vehicle as an old, antiquated, battered and beaten antique... We should study the reasons for this belief and correct this perception pro-actively. In our experience, the following facts contributed to this:



1 User interface & user experience

22

Obviously our user interface and user experience is the first reason, as this is the most visible part of our application, especially if we want to expose this functionality to the generation Y (and soon generation Z) consumers.

Although this is the LEAST of our problems, it receives the bulk of our attention, as it is so visible. This focus actually harms us more from a strategic perspective, than we may be prepared to admit. It is also the easiest of our problems to correct, which means most companies who did or do consider modernization, focus on this... to their own detriment, as this fact hides the real issues...



2 Archaic file and field names

23

We believe that the way in which the DB2 database presents itself to the new generation of SQL literate business leaders hurts us much, much more than we realise. We refer here to 6 – 8 character field names and 6 – 10 character file names, that we programmers used to name our database constructs and elements historically. These SQL “savvy” users take one look at our database and immediately develop an intense hate for the system we love, as they cannot relate to our naming conventions and want to deal and interact with the database using familiar terminology, not “techno speak” or “geek speak”. These users (quite often the ones calling the shots) WILL try and get rid of “this remnant from the dark ages” as a matter of priority, unless YOU change the game...



3 Huge structured monolithic programs

24

Having said this, we believe that the fact that causes most of our trouble, is hidden to a large extent from cursory view: the way we used to develop systems in the 1980’s and 1990’s... We used to develop these huge “structured” monolithic programs, as that was the programming model of the time. As a “rule of thumb”, 80% of the lines of code in these monoliths dealt with database (entity) relationships and database validations. Additionally, there was little to no separation of function, which made maintenance increasingly problematic. The way in which this presented itself, was massive maintenance backlogs and frustrated users.

This particular issue is very closely related to the issue of “Technical Debt”, which we will analyse just now.



Maintenance backlog

23

(b. Maintenance backlog and frustrated users)

These monolithic programs allowed little to no separation of function, which made maintenance increasingly problematic. Every single program that manipulated the database had the same 80% of code repeated over and over. This led to an ENORMOUS amount of duplication, but that was the only option in the 1980's and early 1990's.

This particular vexing issue is the biggest single contributor to "Technical Debt", which I will explain in more detail shortly.

The complexity and potential for problems resulted in massive maintenance backlogs and frustrated users.



Slow stagnation

26

As a result, we stagnated and our magnificent platform was hiding the fact...

During this time, the platform kindly hid the fact that we stagnated... But stagnation in the highly competitive business world of today translates into... DEATH!! This is a very uncomfortable fact to face, but only by admitting it can we change, correct it...MANY of the negative perceptions around the platform are attributable to US programmers/developers. We have a choice now: become relevant by upgrading our skills and tools, our use of especially the database engine, separation of function (MVC or multi-tier architecture) and the ILE programming model or... become EXTINCT!



We need to change ... an accident on its way!

27

We need to change

If we do not step up to the plate, accept and shoulder responsibility, and correct this, we have an accident on the way: and you WILL lose the platform to competitive business pressures.

The fact that you are attending modernization sessions is a step in the right direction!



How much value in your heritage systems?

28

When considering change, you need to know if your heritage applications are of value to the business. This is a fundamental question you need to answer as a matter of priority. You need to know if your applications are worth anything to the business, worth enhancing and investing into. If your heritage applications do not provide significant functionality, it does not make sense to modernise them. It then becomes a simple choice of replacement or redevelopment.

(Bear in mind

- *The business success globally of companies using the platform*
- *The competitive advantage encapsulated in the heritage applications is often DIRECTLY responsible for the success of the companies using them*
- *And the Total Cost of Ownership when delivering these systems to users, clients, suppliers and partners.)*



How do you assess the value?

29

How do you assess the value? An important part of this is identifying the critical operational and strategic constraints

Ask yourself the following questions:



Assess value - Performance Potential

30

- Does the application provide known competitive advantage (such as a unique order entry process, special stock allocation and stock management algorithms, etc.)? List these for the applications in your environment.
- Can you analyse your current business processes and see where they can be improved? Will the system POTENTIALLY allow this improvement?
- Can you analyse your service delivery models and determine what inhibits delivering service directly to end users – your clients, your suppliers, your partners?
- What is the VALUE of the data your business have collected since introducing the application?
- How much have you invested in this application, as this may represent a significant IP (Intellectual Property) investment?



Identify operational constraints

31

- Can you analyse and document the cost structure to deliver the applications
- Do you know what the functional fit of your current application is (s)? As a rule of thumb, new off the shelf applications will deliver between 60 – 75% functionality – usually at the lower percentage range.
- Are the constraints you've experienced a lack of functionality or a weakness in delivery? Analyse these scenarios carefully.
- Can you extend the application with relative ease?



Operational constraints = lack of business agility

32

Based on extensive experience and analysis of many installations that have migrated away from the platform, it is clear that OPERATIONAL CONSTRAINTS are largely responsible for user and business frustrations. This equates to a lack of AGILITY and is consistently being singled out as the most significant factor.

What we found particularly interesting to observe is that many companies who retire the platform due to their constrained applications, often regret it, due to lost functionality and a subsequent decrease in availability and reliability which has been compromised resulting from platform changes. IBM i is legendary for system reliability and availability.

More about AGILITY in a few moments... (or review the provided definition earlier in this hand-out).



What are the options?

33

So what options are available to companies who have identified that the functionality offered is competitive, but operational constraints limit them, especially a lack of AGILITY shackle them, and prevent them from being competitive?



Replacement

34

One option is Replacement

Destroy your proud heritage by scrapping it.

Literally “throwing the baby out with the bathwater”, losing the immense investment in IP and especially the collected business data, the MOTHERLODE of value in your business

Should you consider REPLACEMENT, perhaps opportune that you consider the following fact:



REPLACE

±60–70% functional fit at best

& need to **run in parallel** for 3 – 4 years

MODERNIZE

95%+ functional fit

YOUR competitive advantage !

Future of Heritage Applications

Based on extensive experience, should you replace an application that has customized and been in use for decades by you, the functional fit of your application is quite likely in the mid to high 90%, quite likely encapsulating the very essence, the competitive advantage of your company...

Then, consider the likely disruption of your business, the loss of your intellectual property (the “knowledge” of your business processes and competitive advantaged and encapsulated in your heritage software asset), the loss of insight that is contained within your heritage database, as well as the retraining of all your users and staff...

Compare that with a GRADUAL, Low-risk, non-disruptive modernization process, as you perform routine maintenance of your applications.



Enterprise Modernization

33

Enterprise Modernization, which focuses on unlocking and reclaiming the untapped value in your applications and ESPECIALLY the DATA and extends the life of those applications for proven Return on Investment measurements.

Important! This does not mean you have to modernise your entire application, as so many people suggest!

Your imperative objective should be to deliver maximum value to the business at all times. To do this there are two important considerations that will help you determine how to extract maximum value:

- Identify the 20% of application that generates 80% of transactions.
- Then perform a careful analysis of your maintenance or change requests and this will highlight application functions or programs that have higher levels of maintenance and cause most of the frustrations.

Start with these two identified elements and you will achieve the greatest ROI, and in parallel you will gain confidence with the methodology and approach needed to unlock value from your heritage applications.



How do companies find themselves in this situation?

37

Let's go back for a moment and find out why companies find themselves in the position where they consider scrapping their proud heritage:

One of the reasons for this was that IBM was very successful in protecting and insulating us from underlying changes in the hardware and operating system. One of the competitive advantages of the platform has become its "Achilles heel". The platform is so reliable, so forgiving, that applications developed in the 1980's happily keep on running, despite massive advances in the underlying hardware and software constructs.

Any other supplier would have demanded a rewrite of our applications – just ask your .NET programmers what happens when a new release of .NET is shipped. Added to this was the maxim "If it isn't broke, don't fix it" that most programmers and IT Managers believe in, not recognising that most of these applications have been "broken" from an architectural perspective ever since the introduction of the ILE programming model. The platform insulated us from this harsh reality. How it actually presents itself is in a lack of AGILITY.

Additionally, we are confronted with a few uncomfortable truths:

- A lax attitude within our ranks in terms of KAIZEN - a Japanese term for "Continuous Improvement".
- The perceived risk to update our applications
- The perceived potential disruption to service
- The "Fat and happy" syndrome - biding our time until retirement. We owe our users and our companies much more than that...



Business Imperatives & Technical Debt FUNDAMENTAL Decisions

38

It is imperative that you acknowledge based on extensive experience that you can modernize your heritage application for about 15 – 20% of the cost of replacement or redevelopment, whilst removing ALL of the operational constraints. But, let's analyze some of the considerations a little further...

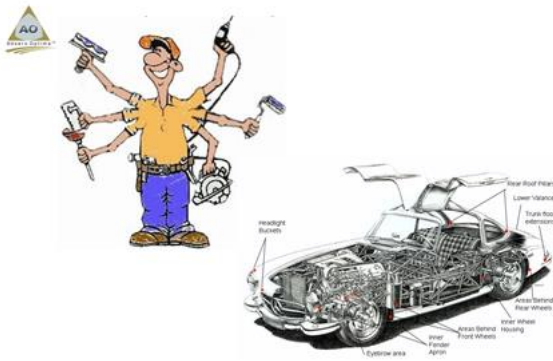


Approach? Tactical?

39

One of the DELIBERATE decisions you will need to make early on, is whether you want to approach modernization as a strategic or tactical project. The fundamental question you need to ask yourself, is how long you would like to benefit from your modernization efforts. If I may, let's use the analogy of upgrading a special sports vehicle that is dear to you...

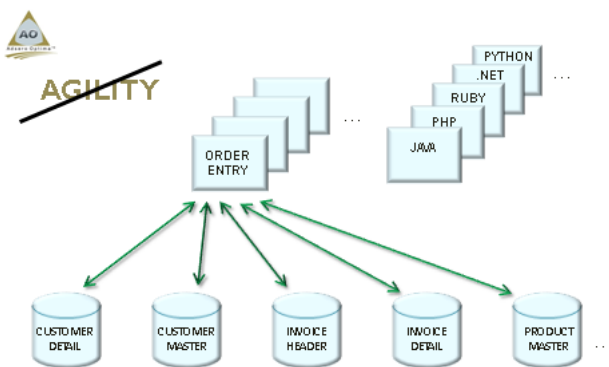
You can improve it on the outside with lots of body putty, special spoilers and fancy paint jobs, but, fundamentally, it is the same old car, with the same old engine and gearbox... OR...



Approach? **STRATEGIC??**

40

You can invest your efforts where it will make the biggest difference and which will last longest: the drive train, the brakes, the instrumentation... whilst also fixing the exterior, but FAR more comprehensively and lasting a lot longer...



Technical debt – why?

To understand technical debt, we have to turn the clock back and consider how most commercial applications originating in the 1980's and early 1990's were developed. This was before DB2 for i (as it is known today) or relational database engines reached maturity and before the introduction of the ILE programming model.

Although we had an integrated relational database engine on the System/38 since announcement, due to the immaturity of relational databases, it was initially effectively a collection of "flat files". All entity relationships and data validation rules (valid values, casting, etc.) was enforced within our HLL code. Should you consider that the average LOB (Line of Business) application from this era consisted out of approximately 1 500 – 2 500 individual program objects, and then consider that EVERY single program that manipulates ANY of the entities (aka physical files or tables), MUST (theoretically) have EXACTLY the same validations rules and relationships enforced, you can begin to understand the significance of ANY maintenance request that involved a change to a field (column) or file (table). To put this in context – MOST maintenance requests will USUALLY entail such changes...

All of us need to acknowledge that we essentially coded relational database logic, using our HLL's of the time, due to the relative immaturity of relational database engines at that time. The 1980's and 1990's was a time of dramatic strategic advances in software engineering in particular.



With the introduction of the ILE programming (OS/400 V2R3) model, shortly followed by the new ILE RPG compiler on the AS/400 in 1993/4 timeframe, all our heritage applications essentially was “broken” from an architectural perspective. One of the fundamental considerations of the ILE programming model, was to facilitate “code re-use”, having a single instance of a piece of logic, and re-using this logic, when required. It also allows for mixing languages dynamically and sharing job control and program storage interchangeably...

DB2 also advanced dramatically in this same timeframe. This was the “golden days” of the AS/400, with rapid development and achievements (CISC to RISC, etc. etc.), especially in terms of sales and excitement around the platform. The remarkable fact (which these days “haunts” us to a certain degree) was how well IBM protected and insulated us from all the changes... our heritage applications happily kept on running, producing the goods, creating the FALSE sense of security that all was well...

Also during this time, our applications moved from essentially online capturing of information with batch processing, to integrated OLTP. It was remarkable how well our applications (and developers) adapted. IBM Rochester achieved some remarkable feats, allowing the developers and users of the AS/400 and successor platforms to achieve something other manufacturers still dream off... However, our applications was ANTIQUATED in their construction methods and was a ticking time bomb...

Add to this the industry move to client/server and then the exposure of our applications to the Internet and mobile devices, using fat or thin clients (or mixture), and we had simply too many demands on our stretched development resources. Many changes became very difficult, especially if it involved field (column) and/or file (table) changes, as we had to find every single instance where that entity (file) or data element (field) was manipulated. If we missed ANY, data corruption potentially occurred...

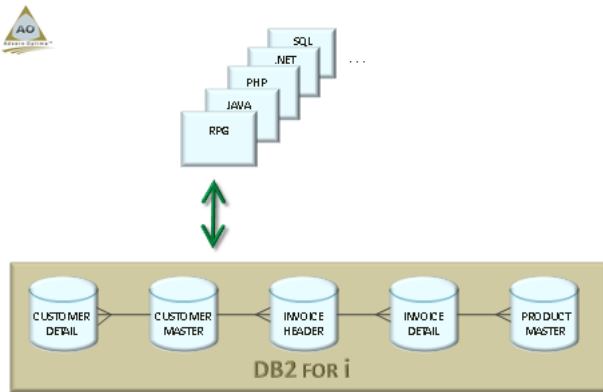
Adding to this highly toxic recipe, many of the early adopters of JAVA and other development environments and languages, accessed and manipulated our DB2 files (tables) directly, DUPLICATING the same validation and relationship code – a recipe for disaster and duplicate/triplicate/quadruple/”to infinite and beyond” maintenance burden. Even worse, the opportunity for data corruption was increased by orders of magnitude...

If it wasn’t for the protection and reliability of this amazing platform and architecture, most companies and developers would have ended up involved in untold horrors...

What should have happened in the mid 1990’s, was the gradual modernization of our applications, leveraging the ILE programming model and the advances in the database engine, implementing “data centricity”. Contrary to popular belief, we actually did have implementations of “data centricity” from the mid-1980’s, with application development tools such as Genesis V and Synon 2E. Synon 2E specifically was decades ahead of its time, from being completely data centric in it’s design and coding paradigm.

However, the squeeze was already on in IT spend, with the C-suite expecting more for less... and our applications was happily producing the goods, so why worry???

As a result, very little REAL modernization happened, at best a little lipstick, if the users complained too much about our “green screens” ...



Data-centric implementation

As indicated in my previous slide, we should have started to adopt and implement “data centrality” in the mid 1990’s, allowing DB2 to look after ALL data validations and entity relationships, making the database “self aware” and placing DB2 (and NOT our HLL code) in control of integrity.

Please note that “data centrality” does NOT mean SQL. And SQL does NOT mean “data centrality”. It means that the database is in CONTROL of enforcing all entity relationships and data validations, preventing ANY client (RPG, JAVA, .NET, SQL, etc.) from injecting ANY inaccurate or incomplete data into your database. Data centrality CAN be achieved with your existing applications, improving your data quality and integrity by orders of magnitude.

ALSO, please acknowledge that we now have SIX (6) “logical” database constructs at our disposal – 3 defined in DDS and three in DDL (DDS View, DDS Join, DDS Multi-Format, DDL EV Index, DDL BR Index, DDL View). The DDL constructs DO NOT replace the DDS constructs. ALL of these constructs are different, each with its unique characteristics. Get to know these INTIMATELY, as the combination of them is an EXCEPTIONALLY powerful tool in your arsenal.

Additionally, it will facilitate an exceptionally AGILE development environment, allowing you to RAPIDLY respond to changes in the business environment. An added benefit is that you can use tried, tested and stable code for enforcing single instances of these rules INSIDE DB2, no longer in our HLL code... And no matter which “client” (RPG, COBOL, SQL, C, C++, JAVA, PHP, Python, choose your language) access our “self-aware” database and receive the same record or result set. It is ALL about single instance, re-usable code. A single place where maintenance is performed, as opposed to n-times...

It also implies that our database will GRADUALLY become a true relational database, with proper normalization enforced, long file and field names exposed, etc...

DUPLICATION & INCONSISTENCIES



Technical debt – Data quality

An issue few people acknowledge is what has happened to our data quality and integrity, due to DECADES of neglect. Again, due to old programming constructs and development methodologies which originated in the 1980's, it is not unusual to see MULTIPLE versions of the truth of the same data element in our business processes, depending where we find ourselves in the business processes. As a result, executives may often look at the incorrect data element when making decisions...

Again this amazing platform SHIELDED us from potential disaster. What should have happened when DB2 on i became the amazing engine it is today, was the implementation of proper data management and data engineering principles. We should have started managing our metadata (the data describing our data) and the quality of our data a LOT more stringently, as is expected on ANY of the competing platforms.

To allow you to grasp the significance of what we are trying to highlight: it is not unusual for us to see more than 50 000 discrete data elements in "use" in the average IBM i LOB application, whilst experience suggests that 1,500 discrete metadata elements would be the number of metadata elements that one could expect to find in the average LOB application.

Now... consider for a moment that your developers are using HLL code to keep the contents of these metadata elements in sync, and you will acknowledge the IMMENSE waste of time that is taking place. Not only that, but consider for a moment the impact of not keeping these elements in sync, or even worse, inconsistencies... Imagine what the quality of decision making is... then, extrapolate this with the advent of predictive analytics and cognitive computing... a potential recipe for a very nice mess...

It is therefore CRUCIAL that you introduce proper Strategic Data Management/Master Data Management principles as a matter of priority, or you WILL drown in data, especially with BIG DATA and cognitive computing approaching all of us at breakneck speed...

So, what needs to happen??



Legacy (heritage) Systems

44

To find out, we need to study briefly how our heritage systems were developed and how the underlying architecture of those applications was composed.

In the past, we used to develop our applications in such a way that all application logic, including all database logic (known as relationships and validations), all user interaction and discrete business logic were encapsulated in large monolithic programs. For instance, it was not strange for mainline transaction processing programs to consist of tens of thousands of lines of code - a programmer's ultimate nightmare.

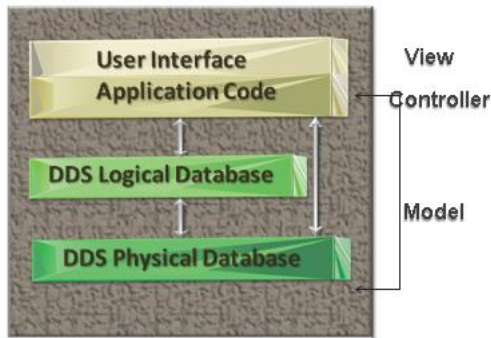
By simply driving database relationships and validations out of the code into the database engine, the lines of code can easily be reduced on average by 80%. The benefits of this, including performance are dramatic – see various resources on our website and ask for case studies.

(for instance, just by re-architecting the application and using DB2 properly at Momentum Life, the largest IBM i user in South Africa, they reduced the run time on a long running job from more than 20 hours every month down to 20 minutes.

With knowledge of our history and the best in class technologies available, let's take a look at the steps we need to take to achieve lasting modernization....migrate and update from DDS to DDL.

For lasting modernization, we need to architect a solid foundation. It is our contention that the only lasting modernization approach possible HAS to start at the fundamental database definition level and the underlying Metadata. Any other adjustments or manoeuvres are tactical at best, providing only a brief respite. In the long term, these do not remove the fundamental barriers to a permanent solution.

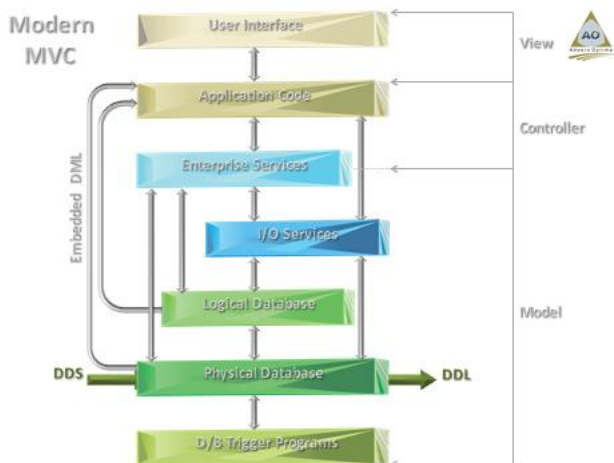
This starts with migrating and upgrading as much as possible of your database from the old DDS definitions to DLL or in other words to DB2 SQL. This alone opens up a considerable amount of value and benefit.



Legacy (heritage) Systems

This is what our monolithic code constructs look like below the covers, where every/most programs handled all the database access, entity relationships, data validations, any business processing and handling all UI/UX, usually in something which after 20 – 30 years of neglect represents a massive POT (not even a plate) of spaghetti of the worst kind... VERY bad for your digestion...

This should over time, at little risk and an no disruption, be changed to this:



ALL entity (file aka table) relationships should GRADUALLY moved out of our heritage code and be enforced by DB2 for i by way of constraints, building upon our NEW DDL constructs. We should then again, in a GRADUAL, non-disruptive process start moving our data validations, out of our code, into DB2 as Triggers. Quite a bit of consolidation of the (duplicate) validation routines will occur at this stage, as you want to end with single instance rules. You will also gradually change the access method to the files, combining both Native IO (also known as Record Level Access or RLA) and result sets. All of this should happen as you perform routine maintenance.

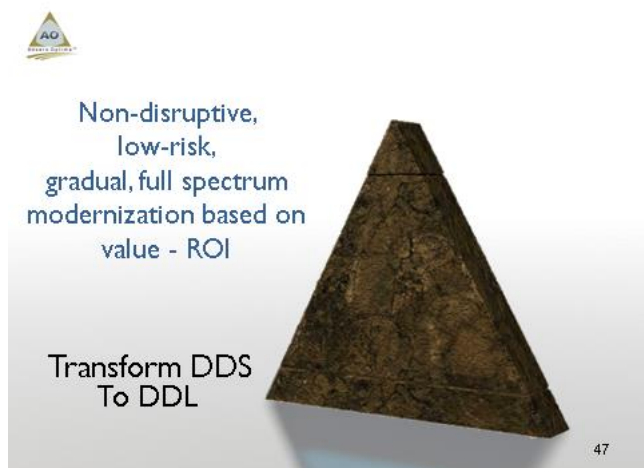
Part of this approach is to gradually separate the three layers or dimensions of our code, into genuine MVC or multi-tier architecture, separating the Model (database) from the View (UI/UX) and the Controller (a very thin layer of code that will EVENTUALLY remain that will orchestrate/conduct) between the database and the UI/UX. The real business unique LOGIC may end up here, depending on the actual design and architectural decisions.

Be very cautious here NOT to step into the trap of “analysis paralysis”. Use the 80-20 rule and your programs that are in perpetual maintenance mode to determine where you want to modernize.

Consistently ask yourself what VALUE (especially from a future maintenance requirement) the business will gain.

ALL of this MUST be underpinned by a central metadata repository (aka Data Dictionary) and formal SDM/MDM (Strategic Data Management/Master Data Management) disciplines. You don't necessarily have to appoint a DBA/DBE, but should certainly assign the responsibility to one of your team members. There usually is a team member that knows the database quite well and has an affinity for data and data integrity.

BUT, do acknowledge that SDM/MDM will only increase in importance with the volume of data that will be thrown at us in future... PREPARE for this!!



The challenge is how to achieve this migration with:
minimum disruption to your users,
the lowest risk,
greatest EASE
as fast as possible and deliver a significant ROI providing a SOLID, long term foundation from which to start extracting maximum benefit from your heritage application.

As I mentioned earlier, it is imperative to get your database definitions into DDL, and move the bulk of your relationships and validations out of your application logic into the database engine. This will provide you with the foundation to start leveraging the incredible capabilities of SQL and the SQE engine. You absolutely want DB2 to do all the "dirty" work for you, allowing you to focus on delivering innovative business solutions and logic.



**Your database is the KEY to
Your FUTURE!!!**

45

The definitions of your database and the data contained within it are the key to unlocking and reclaiming your heritage...

Data-Centricity and using DB2 to the maximum, WILL allow you to ride the “wave” (more like a storm) of BIG DATA (structured and unstructured) with CONFIDENCE...

And UNLEASHING the incredible value that resides in your business data...



**Small, incremental steps,
FAST ROI**

48

A FUNDAMENTAL approach you should adopt from the start, is to introduce small, incremental modernization project steps, producing FAST ROI

Adopt the approach of Modernizing as you maintain.

Trying to run modernization as a separate discrete project, is LIKELY to fail, unless you REALLY have the backing of your C-Suite, due to operational pressures. It is an unfortunate reality and human nature, that under pressure, we revert to what we know best. In the early phases of modernization, CONSISTENCY AND PERSISTENCY will be key, forcing us (developers) to adopt the MODERN way of coding. We CONSCIOUSLY need to break old, BAD habits and it will require an effort.

Our suggestion is that companies take a CONSCIOUS, PUBLISHED decision, with C-Suite support, to allocate x percent (we usually suggest between 10 – 20%) of their time, to modernization and to delay current

projects with that percentage, whatever it is (aim for about 8 hours per week, which is 20% of a 40 hour week). INITIALLY you WILL slow down, but within about a month BOTH your productivity, as well as quality will start improving. After about three months, you will be SIGNIFICANTLY more productive and accurate...

After the initial 3 months concerted effort, you will NEVER look back. Both you and the C-Suite will start to see the benefits and ROI...



No lock-in

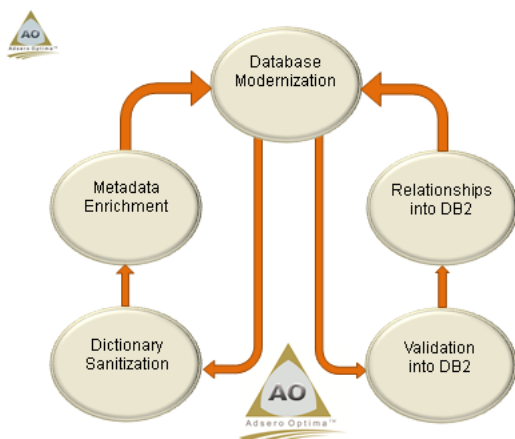
30

Be exceptionally careful NOT to step into the trap of replacing one dependency with another. Stay AWAY from ANY product that “locks you in”. Stay within the IBM toolset (RDi and the compilers of your choice).

If you do make use of tools, ensure that you can use the results (source members and compiled objects) without ANY dependency on the tool.

And please, RPG to JAVA or RPG to >NET “converters” are NOT modernization, unless FUNDAMENTAL re-architecting occurs. Expressing old programming logic in a more “modern” language, does NOT make the result modern...

It fundamentally remains a 20 – 30 year old application...



You should immediately embark on a process of “Incremental Evolution” and perpetual maintenance of your “database”, which is currently probably a selection of “flat files”. In all likelihood, based on extensive experience the past six years, the amount of duplication and inconsistencies in your underlying metadata

will shock you. There is a free tool available, which will very quickly indicate to you what is really going on in your database.

The objective is to achieve this process with no disruption or risk to the business, whilst improving the quality of both your data and metadata by orders of magnitude. It is not unusual to see 90% plus of your metadata elements, the contents of which are currently synchronised programmatically, disappear eventually...

The FINAL objective is a properly normalised database, that is COMPLETELY “self-aware” and “self-enforcing”. This will enable a genuine “data-centric” application environment and complete separation between the database and ANY applications or “clients”.



Gradual clean-up; parallel migration

32

As highlighted earlier on, once you have performed the initial migration/upgrade to the SQL personality of DB2, by moving from DDS to DDL' this forms the foundation to gradually, in a non-disruptive, business-as-usual fashion recover the essence of your system where there remains significant value. Note: We are not recommending that you modernize your entire system. We say “be guided by the value in the business and the operational efficiencies. This should be achieved by doing your standard maintenance, which usually implies that the most maintained parts of your system will be the first to be modernized.



Evolution – not

Revolution!!

33

Gradual code modernisation where VALUE can be unlocked; Facilitate skills transfer and Best Practices implementation

Increase VELOCITY as your team gains confidence.



54

It is FUNDAMENTAL to our suggested approach, that the factors, the components that hurt your business from an AGILITY perspective, be modernized. You should at ALL times ask yourself the question, what BUSINESS benefit you stand to gain by modernizing a particular function. The variables and the measurement will differ from company to company, depending on your industry and competitive pressures as well as your relative positioning.

Focusing and investing your effort initially on the structure and quality of your database, making the database “self-aware”, investing code refactoring on the 20% of business function which usually generates 80% of your transactions, and also those 5 – 10 programs that are in perpetual maintenance mode, your team will quickly develop a “feel” of the benefit that will be gained.

Once you have “picked” these “low hanging fruit”, the focus should then be on continuous improvement and adding advanced functions to our applications...

Keep investing especially on continuous improvement of the database and metadata...



Benefits?

55

As you can see, enterprise modernisation is not a complex or time consuming task. In quite a few cases, the initial move to SQL can be completed relatively quickly, depending on whether you perform this manually or are using a tool. We usually recommend a gradual modernization project, lasting about 2 years. That way you do not place undue demands on your developers, allow them to become completely “au fait” with ILE, RPG IV free, MVC and allows you to focus your attention where it will deliver the biggest benefit to the company.

Our suggested approach usually results in:

- There was no downtime during the migration
- The migration enabled them to leverage DB2 SQL
- There was absolutely NO RISK to the business
- Based on our experience, you can expect a complete modernisation project, on those elements that causes your operational constraints, to cost between 10 – 20% of what a replacement application will cost. Now that is COMPELLING...
- You now have access to a larger and YOUNGER technical resource pool for future maintenance of their systems
- It will re-invigorate your developers and end users
- You will be AGILE, able to respond RAPIDLY to changes in the business environment
- You will have a SOLID foundation to exploit latest advances in Analytics, Social and cognitive computing
- The QUALITY and ACCURACY of your data will improve significantly
- You will likely extend the life of your current heritage application by another 10 years...
- *You will be able to deliver a fast return to value for the business*



SECRET!... WATSON Text Search

**TOP
SECRET**



<http://www.myhofi.com/exphtm/index.htm>
<https://www.linkedin.com/in/rainer-ross-85a594b3/en>

26

What few of you may know, is that you have FREE access to WATSON Text Search capabilities on DB2 for i. We believe that developers should really start to explore the possibilities opened to us by advanced DB2 for i functionality, which will ease the learning curve to exploit the even more powerful OLAP and other WATSON API's functions...

A leading light exploring the more advanced capabilities of DB2 for i, and changing the paradigm about using IBM i and RPG to deliver advanced WEB solutions, is Rainer Ross

(<https://www.linkedin.com/in/rainer-ross-85a594b3/en>) of Germany, with his MyHofi Website:

www.myhofi.com

Do yourself a favour and research and install the OmniFind LICPGM which is FREE with DB2 for i... It will provide you with the perfect opportunity to get your “feet wet” or “fingers dirty”, exploring the immense possibilities advanced analytical functions can add to your heritage applications, and easing into the even more advanced implementations that WATSON opens up...



OLAP and Predictive Analytics in your App



Personality Insights



AlchemyData News



Tradeoff Analytics



Relationship Extraction

57

The API's we believe you should start exploring, doing initially small pilot projects, include the following. Again the objective here is to become familiar with how these API's can be integrated with your LOB applications, and measuring the value and benefit. The indicated API's hold immense promise for commercial OLTP applications in our opinion.

One of the “young Turks” on IBM i, Aaron Bartell, has done quite a bit of work especially around speech recognition and will likely share his knowledge and experience willingly.

Personality Insights

<http://www.ibm.com/smarterplanet/us/en/ibmwatson/developercloud/personality-insights.html>

Personality Insights extracts and analyzes a spectrum of personality attributes to help discover actionable insights about people and entities, and in turn guides end users to highly personalized interactions. The service outputs personality characteristics that are divided into three dimensions: the Big 5, Values, and Needs. While some services are contextually specific depending on the domain model and content, Personality Insights only requires a minimum of 3500+ words of any text.

Relationship Extraction

<http://www.ibm.com/smarterplanet/us/en/ibmwatson/developercloud/relationship-extraction.html>

Unlike general-purpose text analytics tools, Relationship Extraction leverages Watson machine learning technologies. The API can analyze news articles and use statistical modeling to perform linguistic analysis of the input text. It then finds spans of text and clusters them together to form entities, before finally extracting the relationships between them.

Relationship Extraction can be used to build applications such as semantic search engines, business and competitive intelligence tools, knowledge graphs, social media monitoring tools, and more. As the result of

being trained on a news domain, the tool functions optimally with media, ad tech, retail, and online marketing domains.

Can be used in business and competitive intelligence to find information about competitors' products in the news

Useful for brand management; helps find information about your own products

Also assists those interested in research or discovery with finding new topics and changes in frequency

Data Insights

AlchemyData News

<http://www.ibm.com/smarterplanet/us/en/ibmwatson/developercloud/alchemy-data-news.html>

AlchemyData News indexes 250k to 300k English language news and blog articles every day with historical search available for the past 60 days. You can query the News API directly with no need to acquire, enrich and store the data themselves - enabling you to go beyond simple keyword-based searches.

Highly targeted search, time series and counts for trend analysis and pattern mapping, and historical access to news and blog content.

Tradeoff Analytics

<http://www.ibm.com/smarterplanet/us/en/ibmwatson/developercloud/tradeoff-analytics.html>

Tradeoff Analytics is a Watson service that helps people make decisions when balancing multiple objectives. The service uses a mathematical filtering technique called "Pareto Optimization," that enables users to explore tradeoffs when considering multiple criteria for a single decision. When your company makes decisions, how many factors need to be considered? What's the process like? How do you know when you've found the best option? With Tradeoff Analytics, users can avoid lists of endless options and identify the right option by considering multiple objectives.



9 Key Points to Take Home

- Huge value in heritage (legacy) applications
- Modernizing them feasible and realistic
- Do this gradually, business as usual, based on ROI
- Upgrade your skills: RDP(i), ILE, RPG IV Free, XML, HTML
- Move to native SQL engine (DDL & DML) as PRIORITY
- Implement multi-tier architecture (MVC)
- Introduce Strategic Data Management & MDM
- Leverage your heritage data and IP, start a Pilot project to evaluate the use of WATSON and OLAP, assisting your users in their decision making processes
- Significant LIFE remains in our heritage applications

58

9 Key Points to Take Home

1. Huge value in heritage (legacy) applications
2. Modernizing them feasible and realistic
3. Do this gradually, business as usual, based on ROI
4. Upgrade your skills: RDP(i), ILE, RPG IV Free, XML, HTML
5. Move to native SQL engine (DDL & DML) as PRIORITY
6. Implement multi-tier architecture (MVC)

Page 53 of 54

7. Introduce Strategic Data Management & MDM
8. Leverage your heritage data and IP, start a Pilot project to evaluate the use of WATSON and OLAP, assisting your users in their decision making processes
9. Significant LIFE remains in our heritage applications

Where to Find More Information

- IBM White Paper: DDS and SQL - The Winning Combination for DB2 for i - September 2012 Update.pdf
- IBM i Strategy and Roadmap - April 2016 Update.pdf
- IBM Systems Magazine : Special Report: Enterprise Modernization - October 2014
- SG24-8185-00 Modernize IBM i Applications from the Database up to the User Interface and Everything in Between Redbook
- REDP-5095-00 - IBM RedPiece - Tools and Solutions for Modernizing Your IBM i Applications
- AO Whitepaper - Essential Guide to IBM i Modernization Global - V4
- The standard AO Modernization Roadmap customized to individual installation requirements for maximum ROI

- All these and many other available at: www.adsero-optima.com



59



As IBM i users we have a bright future!

60



Marinus van Sandwyk

Email : mbogo@tembotechlab.com
 WEB : www.adsero-optima.com
 LinkedIn : www.linkedin.com/in/mbogo
 Skype : Mbogo33
 Twitter : @Mbogo333

