The Business Case for Automating Data Management

Mark Hobart

Infoboss Limited, UK
Email: mhobart@infoboss.co.uk

Abstract

Back in 2016, IDC forecast a ten-fold increase on business data processed globally by 2025. Combine this with increasing regulatory compliance and demand for quality data, then it is perhaps not a surprise to find that many businesses are feeling somewhat over-whelmed by the enormity of the data management and governance task they face. But, don't despair, there is a compelling business case, both in terms of efficiency gains and tangible financial savings in adopting a strategy for centralising information governance and automating the management of your data.

This paper will explore the emerging problem domain of data and look at the business case for use of automated data management tools to better prepare your organisation for the tsunami of data coming its way.

Keywords: Data management, compliance, data quality management, data governance, unstructured data, business process automation, automated data management.

1 Introduction

It is widely accepted in the data professional community that getting a data management (or data governance) business case off the ground is... Let's just say "challenging"

So, what can you do to enhance, your business case, to get it more readily accepted?

We'll start by looking at what makes a good business case and provide tips to help turbo charge yours.

We'll start by exploring the concept of automation and how applying it to data management and governance can really support your business case and help you to more easily align with other business initiatives.

We'll also look at data growth, the cost of compliance and poor-quality data on the organisation and how you can leverage these important business imperatives to monetise your data management business case.

2 The Information Age

Definition:

The Information Age is the idea that ***access to and the control of information*** is the defining characteristic of this current era in human civilization. (Wikipedia)

We are living in the information age, it's now. But what is fascinating about this definition is the emphasis on the words "**access to and the control of information**".

How many of us can truthfully say that our business has access to and control of its information?

Whatever model you are using to measure your business' current level of data maturity, it's likely that there is significant work to do to get to where you need or want to be.

In recent years, much has been written about the future of business data:

- Ten-fold growth in data by 2025 (IDC);
- Increasing prevalence of unstructured data (the most difficult to access and control) (IDC);
- Increasing levels of regulatory (data related) compliance (GDPR, CCPA etc);
- Increasing cost to business of poor-quality data (Gartner); and
- The rise of automation within business.

What is fascinating about all of these predictions is that they were not only true, but have had a much more profound impact on business than had been originally envisaged.

Given this challenging data environment, our mission as data professionals is to make access and control of information a reality for our businesses in the information age and the first step is a good business case.

3 The Business Case

A good business case must explain the problem and present all possible solutions (including the "do nothing" option). It should be flexible enough to withstand changes in scope and timescale and ultimately enable decision makers to decide upon the best course of action for the business.

It should be aligned with the organisation's strategic goals, have a return on investment (the bigger the better), detail all of the benefits both tangible and intangible and all of the costs.

It should include a detailed risk register and the impact (cost and consequence) of the crystallisation of each risk. It should conclude with a recommendation, i.e. what is your preferred approach? After all, you know your material best, what do you think the business should do?

A cautionary note when compiling a business case for data management and governance is not to fall into the trap of assuming that "Data common sense" is inherent in your audience, it isn't! Many business cases falter or fail because the benefits for good data management can be largely intangible. Your passion for the subject will not win over stakeholders on its own, no matter how passionately you argue the case. If you do not include sufficient tangible benefits and facts, to support your arguments it will not win through and gain approval.

3.1 The Vision

It's worth us taking time to outline the vision and outcomes of best practice data management. To set the scene of what we are trying to achieve.

A leading data professional in the UK, Alex Leigh recently said, "A well-managed, agile and respected data asset is a foundation for the success of any organisation"

A vision that many thought leaders in the data management profession would happily subscribe to.

There are many business outcomes that good data management and governance support. An organisation on top of its "data game":

- **Is delivering on its strategic goals**
- **Has fully operationalised its data management and governance**
- **Has quality and compliant data**
- **Is truly data-driven**
- **Has a data centric culture that is getting better over time**
- **Is able to adapt to changes in the future data environment**

The premise and understanding of this paper is that every business wants to achieve the above for their business.

4 Automation

"Automation is the technology by which a process or procedure is performed with minimal human assistance." (Wikipedia)

Here we have a definition of "automation". It is often used in the context of robots and you might hear the term "robotic process automation" (RPA) especially in the manufacturing sector, but it's not just robots, any process where technology, including computer programs and algorithms, that can help to carry out a hitherto manual process with little or no human involvement can be considered as automation.

Automation is becoming more widespread and the World Economic forum says "that automation will displace 75 million jobs, but create 133 million new ones". These will be jobs that are better adapted to the new division of labour between machines, people and algorithms.

Recent adoption stats from another organisationCentrip indicate that, "3/4 of midsize businesses are using automation in at least one area of their business". They are largely doing this to compete with bigger counterparts, to "punch above their weight".

It's a trend that is unlikely to change and in short, it is going to be on the agenda of every business in the future if it isn't already!

And there are very good reasons why, as there are lots of benefits to automation, with probably the ones associated with saving money and increasing competitiveness being high on the priority list for many seeking to adopt it.

Benefits such as:

- **Lower operating costs;**
- **Faster return on investment (ROI);**
- **Do more with less – maximiselabourresources;**
- **Be more competitive (especially against larger firms);**
- **Operate 24x7;**
- **Increased versatility;**
- **No staff overheads – desk, chairs, HR; and**
- **Automate the boring stuff!**

However, automation is not often (if at all) used in the context of data management where people based strategies and solutions have often, historically, been the order of the day.

However, given the challenges already outlined with data growth, increasing compliance etc... it is a strategy and approach that simply has to change if data management and governance is to be successfully implemented within the organisation.

"Companies will need to increase automation and streamline their organisation if they are not to be overwhelmed by the challenge of sustaining GDPR compliance over the long term." (McKinsey)

Leading management consultancy McKinsey conducted research last year, just after the GDPR deadline. In their recommendations, they made a strong case that automation of data management and governance processes especially around data protection compliance should be considered a necessity going forwards.

One of the reasons for this conclusion was that respondents had observed a tendency to adopt people-based solutions to data management and governance that are simply not sustainable given the predicted growth in data, increasing complexity of compliance and demands for high-quality data within the business.

In short, the age of people-based data management and governance solutions is no longer viable.

So, is automating data management the innovation that data professionals have been looking for?

4.1 Getting the Solution Balance Right

Any data management initiative involves people, process and technology. It's imperative you get the balance right throughout your initiative and build in
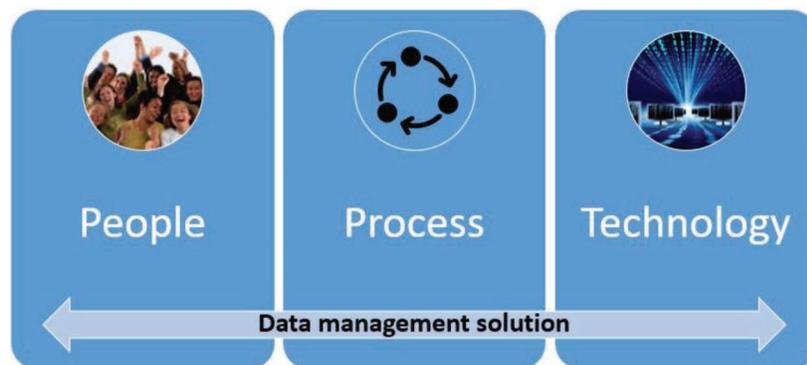

Figure 1   Data management solution elements.

sufficient flexibility in your approach to be able to cope with the potential challenges ahead.

Many initiatives falter or fail completely because they have not set off with clear objectives nor a fully resourced plan of how they are going to achieve the expected outcomes. Too much or too little of any of these elements will set your project on a course for failure very quickly. You must build confidence in your solution strategy early (look for the quick wins) and be ready to constantly innovate to support the business as your strategy and the data landscape unfolds.

I.e. you must get the balance of people, process change and technology right, each one complementing and supporting each other as you go.

A good business case will see a ROI, the quicker the better. However, it can quickly go awry if the solution element blend goes out of sync.

For example, too little (or no) process change and not enough technology can lead to a people based approach that as we have seen and undoubtedly experienced will not be able to cope with the demands of the business now nor in the future.

Too much technology can be expensive and too much process change can severely impact the performance and culture. There is no rule of thumb for success as every business data management initiative will be different, but simply relying on an increase in people-based resources and/or a change in process is not going to cut it in the longer term. You simply must bring technology to the party, giving your data people the tools to do their job properly.

4.2 Automating Data Management

Data management and governance has traditionally involved a lot of manual processes. Checking data, finding information, monitoring policies to ensure they're being enforced and servicing compliance requirements such as those of the data subject to name but a few.

Many of these processes are ad-hoc, mostly labour intensive and I posit that they barely scratch the surface of what is actually needed to do the job properly.

One of the challenges faced by data professionals is being able to access the organisation's data in a way that makes sense to them. Exposes data as they would like to see it and ultimately makes their job easier to do.

Imagine a system that enables you and your data stewards' access to data that has been automatically classified and organised for you, is searchable and easy to explore, analyse and manage. In short, a system that is there to help support you in your job as a data manager or data governance professional and flexible enough to adapt to how you would like to operationalise your data management processes?

A system that is constantly monitoring incoming or changed data to ensure it satisfies compliance and quality requirements.

A system that does not depend on people to do these checks and can do it every day of the year and never complains even when asked to do the boring stuff!

This is what I mean by automating data management. A system that you can train and adapt to your ways of working and help to re-balance the people, process and technology equation to ensure successful data management and governance outcomes for the business.

Figure 2 illustrates the approach taken by my own company infoboss, to automate data management. Essentially, automate the processes of collection, organisation, classification enabling search, analysis and discovery of data by data stewards in a way that is meaningful to them. The application of rules and alerts to effectively monitor the data estate for conformance with your data strategy and policies, enabling a 24x7 data management operation across your entire data estate.
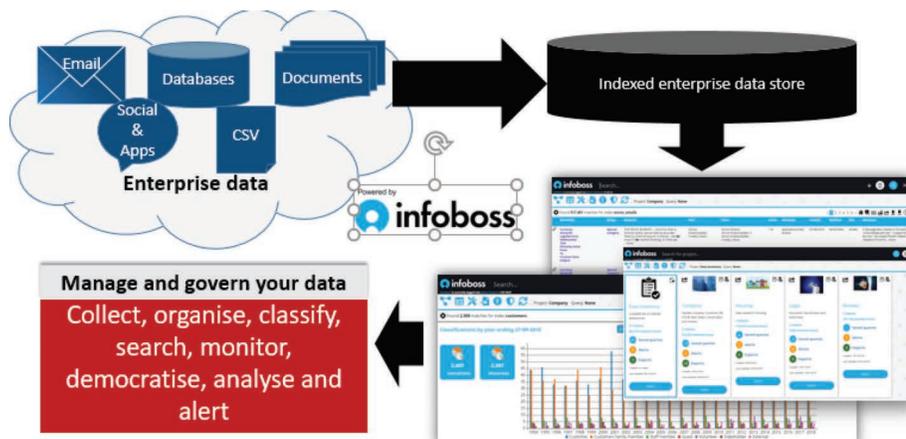

Figure 2   Infoboss automated data management solution.

5 Business Initiative Alignment

Research recently published by Solutions Review, explored the data challenges and opportunities for 2019. They were:

- Customer demand for strong data protection and transparency surrounding the processing of it
- Increasing use of automation, machine learning and artificial intelligence in the business
- Balancing data "attack" and "defence"
    - Agile and exploratory <u>Versus</u> governance and control
- Combining data management and analytics to unlock "big data" insight

I'm certain that most businesses have these items on their data agenda right now. However, in and of themselves they do not represent a sufficiently compelling business case for improving data management and governance.

Automating data management makes alignment with other business initiatives much easier to do and in many cases can be the approach that ensures a successful outcome.

Recognised best practice when building a business case for data management and governance is to look across the organisation for other strategic goals and initiatives that you can align your data management and governance programme with.

That is to say, look at your business strategy and find all goals and initiatives where improved data management will help the organisation to deliver a better outcome for the business.
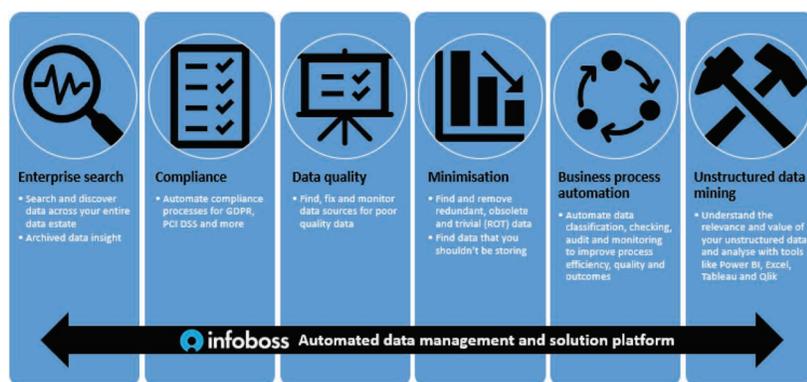
Figure 3   Using automated data management to help alignment with other business initiatives.

Every organisation wants quality and compliant data the consequences of not having it are material (see later)! If these are the topics that the business is leading on, then great. However, it's more likely that they are enablers for other initiatives or to mitigate strategic risks.

Businesses as we have discussed already want to automate processes, they also want to leverage more of the data estate (for example unstructured data) for additional insights using BI tools.

Your business may be on a digital transformation journey or minimising data storage ahead of a cloud migration. Or just after better ways of searching and finding data to keep millennial employees happy! Whatever the "cool stuff" is in your business strategy, explore how you can align with it to get your data management and governance mobilised.

If you invest in an automated data management platform, you are equipping the organisation with tools that can support all of these initiatives and more. And in so doing, dramatically improve the data management and governance of the business. Of course, if the business is not considering doing any of the "cool stuff", how much sweeter would it be for you to announce that you are the one doing it with data!

6 The Data Universe

Back in 2016, IDC predicted a ten-fold increase in data created and replicated by 2025. At time of writing that's just five years away! This will grow the global data estate to circa 163 Zettabytes.

Fascinating though this figure is, they also predicted that 80% of this data will be generated by business compared to just 20% today. A staggering increase on today's data estate size.

6.1 Unstructured Data

In the research findings, IDC also predicted that 80% of the total data estate will consist of unstructured data. It is this prediction that is perhaps the greatest concern for business as historically this has been the most difficult set of data to access, manage and control.

Unstructured data by its very nature, has historically depended upon people to interpret it. The scale of growth in data and the dependency on our ability to manage it better means that people-based processes are going to be put under considerable strain, dare I say no longer sustainable when considering the processing of this data type. It potentially presents the greatest

threat to business compliance efficiency and quality management but a unique opportunity through the use of automation, minimisation and derivation of insight to enhance business performance and generate competitive advantage.

Structured data can be easily accessed and controlled using databases. Unstructured data on the other hand is far more difficult to manage, access and analyse.

6.1.1 Where is unstructured data coming from?

Traditionally this has been office productivity suite documents, PDFs, emails, presentations, spreadsheets and more. Increasingly social media feeds, video, CCTV, scanned documents, photographs and voice data have been joining the data estate. We're also seeing work-force collaboration tools like slack, teams and customer services web chat applications that are adding to the mix. These can be a challenge to monitor and manage especially for compliance purposes.

But perhaps the most significant change is with the Internet of Things (IoT). I.e. data derived from devices and sensors streaming into the organisation. Our ability to collect, store, monitor and make sense of this will be an ever-increasing challenge and dare I say, our greatest opportunity.

IDC produced some research results that suggested that 37% of our digital universe contains data of value and I guess by implication 63% of it has no value! As the data estate fills up with unstructured data how are we going to make that assessment? Are we really going to get a person to look at every piece of unstructured data?

6.2 Take Control of Unstructured Data

There is opportunity to enhance the business case for data management by capitalising upon the IDC research and removing redundant, obsolete and trivial (ROT) data. IDC claim that 63% of data has no value to the organisation. If true, and 80% as we have been advised is unstructured, there is a tremendous opportunity to remove this low or no value data from the data estate. To do this, we need to classify, categorise and understand our unstructured data. Automating data management just got interesting...

Storagecraft.com state that the storage costs per month of one terabyte (TB) of data is £229. So, for an organisation with just 10TB the monthly cost of storage is £2,290 or £27,480 p.a. If 63% of this data is of no value then the company is committing to an annual spend of £17,316 for apparently no value to the business. Remember that this is just the storage costs!

But, if the IDC growth figures are correct and this data grows ten-fold by 2025, by not sorting out the "no value" unstructured data now, you are effectively committing to a future spend of £173,316 p.a. for no benefit to the organisation.

6.3 Relevance and Value From Unstructured Data

Utilising our automated data management approach further, we can seek to tackle the other 37% of unstructured data that has meaning and value to the organisation, by mining it to establish its value and relevance. Using automated data management tools, we can derive hidden insights in the data, automatically classify, categorise, sort and feed derived structured meta data to business intelligence tools to undertake analysis and generate insight from it.

Imagine processing in-bound post as it arrives and looking for certain words or phrases that identify an opportunity or the propensity for fraud before a human has had chance to process it. An automated data management system can process thousands of documents an hour whereas a human may only be able to process a fraction of this amount.

7 Data Protection Regulation

In May 2018, the General Data Protection Regulation (GDPR) became law across Europe and it changed the attitude and approach of businesses within the EU to data. However, it's not just the EU that has been affected. The impact of the GDPR has had a ripple effect globally with the impact of GDPR being acknowledged elsewhere and new data protection regulations being announced and implemented in other countries.

Since the deadline there have been many that believe they're compliant, it's done! However, many believe that this is just the beginning and thought leaders are beginning to reflect upon how they can adopt the GDPR mantra of "Data protection by design and default" effectively operationalising compliance into the DNA of their business and use it to gain a competitive advantage.

Leading up to the deadline, businesses were heavily focussed on what they needed to do to achieve compliance. Ponemon Institute estimated that the average cost per employee of this compliance effort was £2,601. But given my earlier comments about the future state of business data, the effort to date has been viewed by many as the "easy part".

Yet for some, there is a belief that GDPR is done. The reality is somewhat different as many practitioners will tell you that compliance reality is very different. Computer Weekly estimate that 1/3 of European companies are still not GDPR compliant.

McKinsey state, "Companies are resorting to temporary controls and manual processes to ensure compliance until they implement more permanent IT solutions in years to come".

The view of many is that organisations are hopelessly under-cooked with regard to their preparedness and ability to cope with the twin challenges of compliance and massive data growth. The traditional approach has been to throw people resources at the problem, but the odd person and minor process change is not going to help unless the fundamentals of data management at scale are not tackled.

McKinsey also said, "Companies will need to increase automation and streamline their organisation if they are not to be overwhelmed by the challenge of sustaining GDPR compliance over the long term." Thought leaders are:

- Thinking about their operating model, and how they're going to ensure that they will continue to be compliant with the GDPR;
- Considering business processes surrounding an individual's rights?
    - to be informed
    - of access
    - to rectification
    - to erasure
    - to restrict processing
    - to data portability
    - to object
    - to automated decision making and profiling
- Considering how to enforce data retention and privacy policies;
- Anticipating more data subject access and FOI requests;
    - Considering how they are going to find the information to service them?
- Considering data flows in and out of the organisation;
    - How will you monitor and manage them?
    - How will you service a data portability request?
- Considering how they will understand and report a data breach within 72 hours.

Essentially, thought leaders are looking at how to operationalise compliance, streamline their processes. Automating data management and compliance processes is one way of innovating in this area.

7.1　Savings Through Compliance Automation

There is a compelling business case for compliance automation. You won't be able to use technology for every aspect of compliance, but you can certainly use it to do the donkey work for you.

Ponemon institute are well regarded for their research work around data and compliance. Recent studies by them have revealed the true costs of compliance and also for non-compliance which can be almost 2.71 times more expensive!

An alternative way of looking at the cost of compliance is as a percentage of turnover that Duff & Phelps concluded for financial services companies as being 10% of turnover.

Whatever it is in your organisation, then what the Ponemon Institute further discovered through their research was that 29% of the compliance cost could be saved by adopting a centralised information governance strategy and increasing data audits above 5 p.a. Music to the ears of the data professional preparing the business case for investment.

But just imagine if you automated the audit process and did it across all of your data estate all of the time. How much more would you be able to save?

7.2　Cost of a Data Breach

Is there any organisation that does not have data breach as a corporate risk? I very much doubt it.

Again, it's useful to put a monetary value to it. The cost of a data breach (according to IBM and Ponemon) in the UK is £99 per record. So, a leak of 10,000 customer records could result in almost £1m of costs. And that does not include loss of stock value which on average drops by circa 7% in the wake of a data breach announcement.

Yet so little is being done to mitigate this risk. Given the most likely source of a breach will be of data that is not known to the business (possibly held within unstructured data) or downloads of data from other systems, it is crucial that we take control of the data asset and understand the distribution and location of personal and sensitive data across our data estate and who has access to it.

Using a data automation tool like Infoboss, it would not take long to get insight into where your personal and sensitive data is and crucially who has access to it. You can then commence the lock down process to mitigate the risk. Of course, you could wait five years and it will potentially be up to 10 times more of a challenge!

7.3 Servicing Data Subject Access Requests (DSARs)

DSARs and freedom of information (FOI) requests are real challenges for some. The biggest challenge is finding the data in the first place. The regulations specify that you have 30 days to respond. However, for one organisation I worked with recently it was taking 21 days to run a scripted search across data where they believed customer data may be present. Leaving just 9 days to administer, pull it all together, redact and respond. It's only a matter of time before they fail to turn a request around or the volume increases such that they simply don't have the resources to cope.

Another organisation was taking 2 weeks per request. They were getting two requests a month and were happy that the one person they had allocated to servicing DSARs was hitting the turnaround target. But what happened next, they had not planned for. A 'no win, no fee' lawyer moved into town and started hounding their customers for information. Requests quadrupled overnight. Now 4 staff were drafted in, struggling to cope and they had to report a non-compliance to the regulator.

In short, you simply have to consider the operational impacts of compliance and exercise your delivery processes. How fast can you turn around a DSAR? How will you collect the information to report a data breach? What if the volume of DSARs increases? etc, etc . . . Implementing an automated data management solution helps to solve these problems by streamlining the data discovery and collation processes, greatly improving the likelihood of servicing requests for information within the regulatory tolerances and increasing the chances of successfully finding all of the relevant information.

8 Data Quality

Data quality and the costs of not having it, traditionally forms a significant element of a data management and governance business case justification. Rightly so!

These costs come as people accommodate bad data by correcting errors, seeking confirmation in other sources, and dealing with the inevitable mistakes that follow. The result is that without good data management and governance practices, they turn their businesses into highly inefficient data fix factories.

Research by Gartner every year is showing a very worrying trend in this area.

"The average cost to an organisation of poor-quality data is $15 million p.a.!"

This figure was up from $9.7m in the previous year. Staggering growth and a trend that I believe is not going to abate.

The reason why, is that there is an increasing drive within business to look at initiatives to innovate, drive competitiveness and increase business value such as process automation, Artificial Intelligence, digital transformation and predictive analytics. They all rely, nay depend on high-quality data. It is imperative that Data Quality Management becomes a business as usual activity and not a point in time knee jerk reaction to unfolding events!

Quite simply you cannot hope to achieve your business goals if you don't have quality data.

If you feel the Gartner figure is too spectacularly high for your business, then other research by Experian, Price and others estimates that "the cost of bad data to be 15–25% of revenue for most companies". This should enable you to apply a rule of thumb and estimate the costs of poor-quality data on your own business.

Further support (if needed) is that IDC state that "Data workers waste 44% of their time being unsuccessful in their work". How many data people do you employ? Given the potential demand to recruit more data workers over the next few years do you really want to only get 56% efficiency out of them?

9 Conclusion

Data is growing unabated a trend that is unlikely to change. The most difficult type of data to manage "unstructured data" in particular is growing at exponential rates and will soon account for 80% or more of our data estates.

There is increasing regulatory compliance pressure on businesses through GDPR and other regulations.

Businesses are wanting to do more with less and their ability to do so, is increasingly dependent upon initiatives and technology that depends on high-quality data.

Put simply, the competing pressures of data growth, increasing compliance and demand for quality data are putting traditional people based data management and governance processes under immense pressure.

Automating data management provides a cost effective and innovative solution that enables the business to achieve:

- Efficiency gains;
- Financial savings;
- Improvement in data quality;
- Future proof their data management and governance capability;
- Operationalise compliance deliver 'data protection by design and default';
- A competitive advantage; and
- Their strategic goals and initiatives that depend on quality and compliant data.

You can no longer rely solely on people to deliver your organisation's data strategy, data management and governance. It won't work.

There's a compelling business case for automating data management and in so doing enable the "cool stuff" to happen and your business to be better prepared for the tsunami of data coming its way.

Biography

Mark Hobart received his B.Sc. in Mathematics and Computer Science from Lancaster University in 1984. A serial entrepreneur with over 30 years' experience in the computer software industry across a number of sectors. Mark has a passion for innovation and solving business problems with technology. His claim to fame was that he invented and built the first ever online motor

insurance quotation service back in 1998 and met Bill Gates the following year! He has built and sold a number of computer software businesses over the years.

Data is his current focus and passion. To this end, he co-founded Infoboss, an innovative, search-based, automated data management platform.