

January 2020

RESEARCH PAPER

Differentiate Your Product With Unstructured Data Analytic Capabilities

Sponsored by



CONTENTS

•	Introduction	р3
•	Key findings	p4
•	Unstructured data - how much is there and why does it matter?	р4
•	Obstacles to insight	p8
•	Unstructured data's structural challenges	p9
•	New ambitions, old technology	p10
•	Conclusion	p13
•	About the sponsor, OpenText	p14

This document is property of Incisive Media. Reproduction and distribution of this publication in any form without prior written permission is forbidden.

Introduction

The digital economy is generating unprecedented quantities of data. Businesses have more data at their disposal than ever before. Some of the most valuable data is unstructured and many businesses are struggling to unlock its full potential. Without it, organisations aren't generating a complete picture of their operations. Software providers and channel partners can help their customers to extract insight – and subsequently value – from this mountain of unstructured data, and differentiate themselves from their competitors.

Computing surveyed 150 individuals actively involved in using, testing, evaluating or procuring data analytics tools at their organisation. They represent companies from a wide variety of industries, including banking and finance, logistics, manufacturing, retail and education to explore the state of data analytics in end user organisations, opinions and use cases around structured versus unstructured data, the obstacles many organisations are facing in extracting value and insights from their unstructured data and how software providers and channel partners can respond to these challenges, add value and become key players in the shaping of data analytics strategy and software.

Key findings

- Sixty percent of respondents stated that a majority between 51 and 100 percent of their data – was unstructured.
- The top motivators for unstructured data analysis were gaining process efficiencies, including automation, and the desire for greater customer insight.
- The top three use cases for unstructured data analysis were all very much customer focused trend tracking, sentiment analysis and business communication analysis.
- Despite the competitive benefits that unstructured data analysis can confer, only 50 percent of those we surveyed were actually doing it.
- Sixty-nine percent of respondents said their organisations found it quite or extremely difficult to gather and analyse unstructured data.
- Fewer than a quarter of respondents said their organisation had been successful in gathering and analysing unstructured data.
- The primary obstacles to deriving insight from unstructured data were cost of implementation, a shortage of technical expertise, and the technology itself falling short and being difficult to integrate.
- Only 36 percent of respondents were employed in organisations where access to data analysis was democratized with the empowerment of individual users.
- There is a mismatch between reported levels of data democratization (36 percent) and the proportion of respondents reporting the use of data visualisation dashboards which are supposed to enable this democratization (53 percent).
- Only 30 percent of our respondents were using Artificial Intelligence/Machine Learning as part of their data analytics platforms.

Unstructured data – how much is there and why does it matter?

The first step to understanding why unstructured data is so important, is understanding just how much of it exists in relation to more structured data.

Organisations have a torrent of data pouring in each day and it is both increasing in volume and diversifying in format. A few years ago, most unstructured data used to consist of mainly text – emails, text files and so on. A great deal of unstructured data is now in the form of video or photos from social media, audio files, and website content.

There is also a considerable amount of data already in the public domain that can be accessed by business customers – such as government or census data and market data from organisations like the IMF. This public data is often in unstructured or semi-structured formats.

Computing asked respondents to our survey to share with us just how much of their organisation's data is unstructured. We can see from the illustration below that 60 percent of those answering this question said that *a majority of their data* – between 51 and 100 percent – was unstructured. This finding answers the first part of the question of why unstructured data is important. Unstructured data matters partly because it simply comprises a majority of the data available for analysis.

Fig. 1 : What proportion of your organisation's data is unstructured (such as emails and other communications, contracts, documents, and social media posts)?



Unstructured data also matters because of the insight that can be drawn from its analysis. *Computing* asked the respondents whose employers were analysing unstructured data what they hoped to achieve from doing so (Fig. 2, *see next page*). The top motivator, cited by 43 percent, was the scope for reduced spend through greater efficiency/automation. Certainly, the potential for productivity gains by the business process analysis and optimisation that unstructured data analysis can inform is huge. Business processes of all kinds involve increasing volumes and variety of data and more applications are involved in business processes than in the past.

Many businesses are looking to optimise business processes via greater automation in order to become faster and more flexible, and ultimately deliver the speed and personalisation of services that consumers are increasingly demanding. Replacing slow, manual, error-prone processes with automated, integrated workflows can bring about huge efficiency gains. However, in order to be successful, the move to automation must be data-informed to ensure that optimisation is applied where it can deliver the biggest gains.

Fig. 2 : What are the main motivations for your organisation's decision to gather and analyse unstructured data? [three maximum]



Only marginally less a motivator was customer insight, which was cited by 42 percent. Customer service data is typically unstructured – phone calls, email, social media posts, review sites and even chatbot interactions. Gaining visibility of this data and an ability to visualise, slice and dice it enables businesses to gain insights into what their customers want now – and predict what they are likely to want in the future. Businesses can get to know their customers on a level that is impossible when structured data is the only source of information.

In a related question, we also asked survey respondents exactly what they were using their unstructured data for. The top three use cases were all very much focused on customers, as the diagram beneath illustrates. The most popular use case was trend tracking. There are two levels of trend tracking. Many of our respondents will be analysing historical unstructured data to look for trends in their customer behaviour which they can utilise in the future, perhaps by segmenting their customers and targeting them accordingly.

Fig. 3 : What does your organisation currently use unstructured data analysis for? [select all that apply]



Using AI-based text mining tools allows businesses not to just see what has occurred but to mine huge datasets from many sources – blogs, social media, review sites are just a few examples – to make predictions about market trends for key topics, products or services. It is interesting that the second and third most popular use cases of unstructured data analysis were customer sentiment tracking and business communications analysis. Sentiment tracking is a machine learning-led application of AI to predict what customers will want and deliver against expectations it by means of hyper personalised services and recommendations.

Less loyal customers can be identified and targeted, which means that advertising can be delivered with laser-like accuracy. Business Communications Analysis also focuses on customers to ensure that they are happy with the type of communications they have with companies rather than being dissatisfied with interactions they deem impersonal or faceless.

Obstacles to insight

The insight that unstructured data analysis can provide is unparalleled. The use cases set out above illustrate this potential. Given these benefits it is reasonable to expect that the majority of respondents would be analysing this data. The problem is, as revealed by this research, that 50 percent of them aren't.

This poses an obvious question. Why?

Computing asked a series of questions about the day-to-day technical and organisational aspects of unstructured data analysis and it is the answers to these questions which provide a hint as to why businesses aren't making better use of this precious resource. The first is shown below in Fig. 4. The answers to this question make it immediately apparent that many more organisations than not are finding it either difficult or very difficult to gather and analyse the relevant data.

Fig. 4 : On a scale of 1 to 5, '1' being 'extremely difficult' and '5' being 'extremely easy', how easy would it be for your organisation to gather and analyse its unstructured data at present?



A related question asking respondents to rate the success of their organisations at gathering and analysing unstructured date yielded a result which can probably be best described as mediocre. Twenty-four percent awarded their businesses a moderately or highly successful rating. Eighteen percent did the opposite. The remaining 58 percent chose faint praise and went with a middle rating of neither particularly successful nor unsuccessful.

Unstructured data's structural challenges

What's going wrong? *Computing* asked respondents to choose three main obstacles to their organisations ability to gather and analyse unstructured data. Coming as a surprise to absolutely no one was the most frequently cited issue – cost of implementation (45 percent.) So far this makes sense, but it's worth stepping back to view this from a partner perspective to ask why costs are so steep.

Fig. 5 : What are the main obstacles to your organisation's ability to gather and analyse unstructured data? [three maximum]



Public cloud economics were a beguiling prospect to businesses desperate to find cost savings in the harsh business climate prevailing in the wake of the financial crisis. However, the reality hasn't quite matched expectations because of the complexity that multi-cloud and hybrid cloud infrastructures have engendered.

It's impossible to analyse data if you can't see it, and this growing complexity has limited visibility of data and made it much harder to integrate data and applications. It is not a coincidence that the second most frequently cited obstacle was lack of expertise, which 41 percent of respondents mentioned. The expertise needed to integrate and manage this fragile web of applications and underlying data is in short supply – and that's before you even get to the data science and analysis skills. Seventy-eight percent agreed, albeit to varying degrees, that, "it is difficult to successfully integrate data analytics tools into our technology stack."

These findings represent a real opportunity for channel partners looking to add greater value to fill skills gaps where they exist on a short-term basis. Employing staff full-time on permanent contracts is not necessarily an approach well suited to businesses trying to move to a more agile way of working – which many are trying to do. Working Agile involves small teams to deliver projects which then break and reassemble as required. Ideally you bring people to the project, rather than the other way around. By reducing the costs and risks associated with permanent employment, the use of contract hires from channel partners for project management, consultancy and technical skills at all levels could help businesses to show some quick wins for data analysis projects. It pays to lean on those that have been there and done it before.

New ambitions, old technology

After cost and a lack of expertise, the next most commonly raised obstacles to unstructured data analysis both related to technology. Twenty-seven percent of respondents raised technology falling short and poorly integrated technology as a problem. We have discussed the difficulties of integration already but the subject of technology failing to keep pace with requirements is one that requires further examination.

There are likely to be two dimensions to the technology problem. The first relates to organisational structures – who is responsible for data analysis? The second is more about the tools themselves. These two dimensions are closely related – the easier users find it to visualise and cut data, the more democratised your data analysis can be. Organisational and process barriers are significant. Eighty-four percent of respondents agreed to at least some extent that, "it is difficult to successfully integrate data analytics tools into our business processes."

Of course, most organisations have a mixture of analysis methods being applied. Forty-two percent of our respondent's employers had a dedicated central team and 44 percent had local specialist teams throughout their organisations. In 36 percent of cases, respondents state that they had a more democratised systems where individual employees were empowered.



Fig. 6 : What organisational structures do you have in place to deliver data, insights and business value from data? [select all that apply]

Fig. 6 indicates some overlap with organisations running multiple scenarios. For example, it's quite possible to have a limited number of users analysing their own data alongside a central resource to help with more intricate analysis. The lack of hybrid approaches here shows the lack of maturity in most organisations data analysis strategy – particularly at those larger organisations who stand to benefit from a combination of methods.

The actual tools and technologies in use vary significantly, as Fig. 7 shows (see over).

Fig. 7 : How does your organisation analyse its data? [select all that apply]



The findings to this question provide a series of clues into why unstructured data analysis is providing so little collective insight. The fact that there is a mismatch between reported levels of data democratisation and the proportion of respondents using data visualisation dashboards, which are supposed to enable this democratisation, indicates that the visualisation dashboards aren't as user friendly, or suitable for broad use, as they should be.

The essence of self-service dashboards is that they enable the user to slice and dice data with ease but the combined findings here suggest that they still need considerable input from data experts to get to the insight. This time lag on insight is profoundly unhelpful for businesses trying to make more informed decisions faster – a fact not lost on the 90 percent of respondents who agreed to at least some extent with the statement, "data democratisation is the best approach to building a data-guided organisation."

The high showing of Business Intelligence tools also hints at part of the reason why unstructured data analysis is proving so difficult for so many businesses. BI tools just weren't designed with the huge quantities of unstructured data that businesses need to analyse in mind. Many BI dashboards only identify events after they have occurred. This is not insightful – it effectively amounts to monitoring. It can tell you about an event in the past but there is often limited context around data points, so understanding and insight is equally limited.

Thirty percent of our respondents were using AI or machine learning, and whilst these algorithms can derive insights from huge data sets, they generally require data science skills to operate, and we have already established that these are in very short supply. It is also likely that the large amount of bespoke and third-party tools which our respondents report are creating further complexity and slowing down the process of deriving insight.

Our research suggests that a considerable proportion of businesses who are trying to bring unstructured data into their analysis are doing so by trying to adapt relatively old analysis tools to a very new challenge. The reported success rates suggest this is unlikely to prove a winning strategy.

Conclusion

It has become a common cliché in commerce and technology to assert that data is the new oil. This is a limited analogy because, unlike oil, data is increasing in aggregate volume every second of every day. Data scarcity is not an issue that business customers are struggling with. What they are battling with is how to realise value from the mountain of data available to them – particularly unstructured data. Sixty percent of participants in our research told us that most of their data was unstructured.

What were respondents hoping to gain from unstructured data analysis? Reduced costs via greater efficiencies and automation and customer insight were the top motivators. Interestingly, the most popular use cases were all very much focused on customers – trend tracking, sentiment tracking and business communications analysis. The desire to understand the motivations and thoughts of customers about various products, services or brands is widespread, as is the desire to predict what customers will want.

Unstructured data analysis sounds like an essential capability for businesses keen to digitise their goods and service propositions. However, half of the organisations we spoke to were not doing so. Almost 70 percent said that their organisations found it difficult or extremely difficult to gather and analyse unstructured data and fewer than a quarter said they had been successful in doing so.

The most commonly encountered obstacles to gaining value from unstructured data was the cost of implementation and a lack of technical expertise. These costs and technical difficulties arise, in part, from the head spinning complexity of hybrid cloud infrastructure. This complexity gives rise to several issues but relevant here are the difficulties of data visibility and integration. A huge 78 percent of our respondents agreed that it was difficult to successfully integrate data analytics tools into their technology stacks.

Analytics technology is also falling short. These failures are in part caused by a mismatch between organisational structures and the technology, and, in part, because a lot of the technology was designed for a business environment that was not awash with unstructured data.

Perhaps the most telling finding is that only 30 percent of respondents were using AI or machine learning to analyse their unstructured data. Businesses are looking for help in defining strategies and implementing the right technology to help them extract insight, and subsequent value, from their data mountains.

Software vendors that can help organisations overcome these challenges and tap into the benefits on offer – leveraging the clear hunger in the market for capable solutions – stand to flourish as unstructured data analysis becomes ubiquitous.

Often, the best way to do so is by embedding already available white-label solutions – bypassing the massive investment, time and other resources required to develop AI-based analytics tools. This leaves vendors to focus on what they do best, and allows them to create products that can make unstructured data analysis more democratised. When simple to use but highly capable tools are in the hands of more end users, informing their day-to-day decision making, unstructured data analysis reaches its potential.

83 percent of those who said it is extremely easy for their organisation to gather and analyse its unstructured data, revealed that the analysis of unstructured data had been extremely successful at their company. Software vendors and other partners can help make this a reality for more businesses.

About the sponsor, OpenText

OpenText, The Information Company, enables organizations to gain insight through market leading information management solutions, on-premises or in the cloud.

OpenText provides organisations with key capabilities for managing information at any stage of the information lifecycle, and the OpenText OEM Program makes this same tech available to other technology vendors to be customized, extended, embedded and white-labelled for use in their own products and services.

A key area of focus for the OpenText OEM Program is the analytics space. OpenText's analyse, predict and report solutions provide OEM partners with the ability to readily integrate a host of analytic capabilities, including unstructured data analytics.

For more on the OpenText OEM Program, please visit our website or contact us to get started today. You can also stay up to date on trends and opportunities that tech vendors care about by reading the OpenText OEM blog.

OEM Program: www.opentext.com/products-and-solutions

Contact: www.opentext.com/about/contact-us

Blog: blogs.opentext.com

opentext[™]