



INSTITUTO POLITÉCNICO DE COIMBRA

INSTITUTO SUPERIOR DE CONTABILIDADE E ADMINISTRAÇÃO DE COIMBRA

AUDITORIA EM TEMPOS DE
BIG DATA & ANALYTICS
REQUISITOS MÍNIMOS DE CONTROLO

Vanessa Filipa Gomes Vieira

Orientadora: Doutora Isabel Pedrosa (Professora Adjunta, ISCAC)

Coorientador: Dr. Bruno Horta Soares (Professor Adjunto Convidado, ISCAC, e Presidente do ISACA Lisbon Chapter, Especialista)

COIMBRA

Outubro de 2016



INSTITUTO POLITÉCNICO DE COIMBRA

INSTITUTO SUPERIOR DE CONTABILIDADE E ADMINISTRAÇÃO DE COIMBRA

AUDITORIA EM TEMPOS DE
BIG DATA & ANALYTICS
REQUISITOS MÍNIMOS DE CONTROLO

Vanessa Filipa Gomes Vieira

Dissertação submetida como requisito parcial para a obtenção do grau de
Mestre em Auditoria Empresarial e Pública

Orientadora: Doutora Isabel Pedrosa (Professora Adjunta, ISCAC)

Coorientador: Dr. Bruno Horta Soares (Professor Adjunto Convidado, ISCAC, e Presidente
do ISACA Lisbon Chapter, Especialista)

COIMBRA

Outubro de 2016

*“Ninguém escapa ao sonho de voar,
de ultrapassar os limites do espaço onde nasceu,
de ver novos lugares e novas gentes.
Mas saber ver em cada coisa, em cada pessoa,
aquele algo que a define como especial,
um objecto singular, um amigo, - é fundamental.
Navegar é preciso, reconhecer o valor das coisas e das pessoas,
é mais preciso ainda”
Antoine de Saint-Exupéry*

Agradecimentos

Este espaço é dedicado àqueles que, de alguma forma, contribuíram para que esta dissertação fosse realizada. Não sendo viável nomeá-los a todos, há no entanto alguns a quem não posso deixar de manifestar o meu apreço e agradecimento sincero. À minha orientadora, Professora Doutora Isabel Pedrosa, por poder sempre contar com o seu entusiasmo contagiante, com a sua dedicação e com a sua palavra amiga, de reconhecimento e de incentivo a cada momento, a confiança depositada as quais contribuíram decisivamente para que este trabalho tenha chegado a bom termo. Ao meu Coorientador Doutor Bruno Horta Gaspar pelo apoio, disponibilidade manifestada e a partilha dos seus sábios conhecimentos. A todos os entrevistados que tornaram possível este estudo, pelo grande profissionalismo e tempo precioso despendido a colaborar na entrevista de forma entusiasta. Aos colegas do Mestrado em Auditoria Empresarial e Pública 2014/2016, com quem vivi um ambiente de verdadeira aprendizagem colaborativa. Aos meus amigos, por todas as horas em que tiveram a paciência de me apoiar nesta fase, mas acima de tudo pela força, pelo sorriso e pela disponibilidade com que sempre me ajudaram. Entre elas um agradecimento especial à Patrícia por dar a palavra sempre certa no momento certo, à Cláudia F. pelas horas que me incentivou sempre a continuar e pensar no objetivo final, à Leandra por me ter acompanhado sempre em todas as fases académicas da minha vida, à Mara pela partilha de bons momentos e à Cláudia O. pelo companheirismo. Obrigada pela vossa amizade.

Os últimos são os primeiros e os mais importantes na minha vida: ao Gonçalo, meu namorado, companheiro e sempre amigo, ouvinte atento de algumas dúvidas, inquietações, desânimos e sucessos, pelo apoio, pela confiança, pelo carinho pela valorização sempre tão entusiasta do meu trabalho ao longo deste percurso. Obrigada por todo o amor e dedicação. Ao meu pai, à minha mãe e aos meus irmãos, Ricardo, Carina e Cátia, e à minha afilhada Lara por sempre me incentivarem perante os desafios, a fazer mais e melhor, quero partilhar convosco a alegria de os conseguir vencer continuamente! Entenderam as minhas faltas e momentos de afastamento e reclusão e demonstraram o quanto era importante estudar, mesmo não tendo eles a mesma oportunidade no passado. Uma palavra de reconhecimento muito especial para eles, pelo amor incondicional e pela forma como ao longo de todos estes anos, tão bem, souberam ajudar-me. A eles dedico este triunfo, obrigada por tudo!

Resumo

O estudo debruça-se sobre o conceito de “*Big Data*” e “*Analytics*”, com o objetivo de investigar o seu potencial como ferramenta capaz de agregar valor para as organizações de diversos segmentos, revertendo em estratégias de negócio num contexto de auditoria, numa conjuntura de requisitos mínimos de controlo. Estamos perante o fenómeno *Big Data & Analytics* quando nos referimos ao crescimento do volume e variedade de dados, gerados a uma grande velocidade, agregado também à veracidade e valor da informação em função dos avanços tecnológicos. O grande objetivo é apresentar soluções para dar resposta às novas tendências, à análise metódica dos dados para melhorar as decisões.

Neste trabalho apresentam-se conceitos formulados por organizações e especialistas sobre o *Big Data & Analytics*, uma das tendências organizacionais, e a sua ligação ao trabalho do auditor, nomeadamente nos desafios que representa para estes profissionais. Os resultados do estudo de caso apresentado são o reflexo de várias entrevistas realizadas a profissionais de auditoria a nível nacional. A pertinência desta amostra está relacionada com o acompanhamento que os auditores devem fazer perante a evolução das tecnologias e na realização de investimentos na formação e implementação de novas técnicas e metodologias.

Estamos numa era em que se tornou imprescindível encontrar formas de trabalhar com todo o volume de dados gerados diariamente. Desta forma, existe um conjunto de desafios a ultrapassar no que diz respeito à qualidade de dados, de segurança e de formação de utilizadores, para que analisar dados seja mais rápido e eficaz. Torna-se necessário o estudo das técnicas, das ferramentas e estruturas que suportam o processamento da quantidade de dados, para dar um impulso à *Analytics* aplicado em *Big data* ainda que este seja um caminho longo a percorrer pelas organizações de Portugal.

Palavras-chave: *Analytics*; Auditoria e Tecnologia; *Big Data*; Informação; Organizações.

Abstract

The study focuses on the concept of "*Big Data*" and "*Analytics*", with the objective of investigating its potential as a tool capable of adding value to organizations of different segments, reverting to business strategies in an audit context, in a Minimum control requirements. We are dealing with the *Big Data & Analytics* phenomenon when we refer to the growth of the volume and variety of data, generated at a great speed, also added to the veracity and value of information as a function of technological advances. The main objective is to present solutions to respond to new trends, methodical analysis of data to improve decisions.

In the present work are presented concepts formulated by organizations and experts on *Big Data & Analytics*, one of the organizational trends, and its connection to the work of the auditor, in particular on the challenges that it represents among these professionals. The results of the case study presented are a reflection as consequence of several interviews conducted with auditors in Portugal. The relevance of this sample is linked to the follow-up that auditors must do in the evolution of the technologies and in the investments made in courses and training and on the implementation of new techniques and methodologies.

We are in an age where it has become imperative to find ways to work with all the volume of data generated daily. In this way, auditors are facing a number of challenges to overcome in terms of data quality, security and users' training, to develop information analysis in a faster and more efficient mode. It is necessary to study the techniques, tools and structures that support the processing of big amounts of data, to give an impetus to *Analytics* applied in *Big Data* although there still is a long path to go to through the Portuguese organizations.

Keywords: *Analytics*; *Big Data*; Audit and Technology; Information; Organizations.

Índice Geral

Agradecimentos	ii
Resumo	iii
Abstract	iv
Índice Geral	v
Índice de Tabelas	viii
Índice de Figuras	ix
Lista de Acrónimos e Siglas	x
1 Introdução	1
1.1 Motivação e pertinência	3
1.2 Objetivos do trabalho	3
1.3 Principais contributos deste trabalho	4
1.4 Estrutura do relatório	4
2 Estado da Arte	6
2.1 <i>Big Data</i>	6
2.1.1 A realidade do processamento de informação	10
2.1.2 O que é o Big Data	12
2.1.3 Atributos do <i>Big Data</i>	17
2.1.4 Por detrás de uma grande decisão está sempre uma grande pergunta	21
2.2 <i>Analytics</i>	22
2.2.1 Âmbito de aplicação em Auditoria	23
2.2.2 A Transformação de Dados	24

2.2.3	Vantagem competitiva	25
2.3	Auditoria - Requisitos mínimos de controlo	27
2.3.1	Riscos do <i>Big Data</i>	31
2.3.2	COSO – Controlo Interno.....	38
2.3.3	Como <i>Big Data</i> e <i>Analytics</i> influencia a Auditoria	45
3	Metodologia	48
3.1	Método “Estudo de Caso”	48
3.2	Estratégia da investigação.....	48
4	Estudo de Caso.....	49
4.1	Contexto da pesquisa	49
4.2	Desenho da entrevista	49
4.3	Análise de resultados	51
4.3.1	Amostra do Estudo de Caso	51
4.3.2	Controlo dos dados gerados e responsáveis.....	52
4.3.3	Responsáveis pela <i>Analytics</i>	53
4.3.4	Ferramentas de Extração de Dados	53
4.3.5	Técnicas concretas para extrair, recolher, tratar ou analisar os dados	54
4.3.6	Gestão de risco e privacidade dos dados.....	54
4.3.7	Estratégia de Negócio Versus Estratégia de <i>Analytics</i>	55
4.3.8	Objetivos relacionados com o <i>Big Data</i> e/ou <i>Analytics</i>	56
4.3.9	Recursos Humanos versus <i>Analytics</i>	56
4.3.10	Orçamento versus <i>Analytics</i>	56
4.3.11	Investimento em <i>Software</i> de <i>Analytics</i>	56

4.3.12	Desafios no tratamento de <i>Analytics</i>	57
4.3.13	<i>Big Data</i> Versus Auditoria	57
4.3.14	<i>Big Data</i> a tendência do futuro	58
4.3.15	Benefícios do <i>Big Data</i>	58
4.3.16	<i>Big Data</i> versus Riscos	59
4.4	Discussão dos Resultados	59
5	Conclusões e Trabalho Futuro	63
5.1	Trabalho futuro	65
5.2	Limitações do estudo	66
	Referências bibliográficas	68
	Anexo A. Artigo publicado 11ª CISTI'2016 – Conferência Ibérica de Sistemas e Tecnologia de Informação	74
	Anexo B. Entrevista a Auditores	82

Índice de Tabelas

Tabela 1: Definições para o termo <i>Big Data</i>	14
Tabela 2: Amostra do Estudo de Caso, editada pela Autora.	52

Índice de Figuras

Figura 1: Dados Não Estruturados - Uma possibilidade de desenvolvimento social.	6
Figura 2: Visão Geral <i>Big Data</i> .	8
Figura 3: Dos dados ao conhecimento. Editado pela autora.	9
Figura 4: O Processo da Informação no contexto da tomada de decisão.	11
Figura 5: Dados Estruturados VS Dados Não Estruturados.	16
Figura 6: Os 5 V's <i>Big Data</i> .	20
Figura 7: Nova abordagem de <i>Big Data</i> .	22
Figura 8: <i>Analytics</i> .	23
Figura 9: Soluções para o problema do <i>Big Data</i> .	26
Figura 10: Riscos associados ao <i>Big Data</i> .	32
Figura 11: Análise de Riscos.	35
Figura 12: Hierarquia das Técnicas Utilizadas na Realização de Procedimentos Substantivos de Revisão analítica. Elaboração Própria	38
Figura 13: <i>Big Data & Analytics</i> .	41
Figura 14: Fontes de dados utilizados na realização de técnicas de <i>Analytics</i> .	42
Figura 15: COSO.	43
Figura 16: Ciclo da utilização da técnica de <i>Analytics</i> no conhecimento do negócio.	44

Lista de Acrónimos e Siglas

CI - Controlo Interno

COSO - Committee of Sponsoring Organizations of the Treadway Commission

EI - "Intelligent Economy

IDC - International Data Corporation

TI – Tecnologias de Informação

TIC - Tecnologias de Informação e Comunicação

Zetta - Zettabytes, símbolo Z é um prefixo do Sistema Internacional de Unidades que indica um fator de 10^{21} , ou 1 000 000 000 000 000 000.

5 Vs - velocidade, volume, variedade, veracidade e valor

1 Introdução

A nossa sociedade dá cada vez mais relevância à informação como sendo uma das suas preocupações mais dominantes de todo e qualquer tipo de organização. Essas preocupações passam por ter uma infraestrutura adequada quer para *Analytics* quer para o armazenamento, apresentando-se como uma prioridade para a gestão eficiente. Este é um cenário típico dos dias de hoje, pelo que, uma das motivações deste trabalho consiste em conhecer de forma mais profunda o cenário da utilização de dados, em particular, no contexto *Big Data* e perceber o impacto que tal poderá ter na forma como os auditores realizam as suas tarefas.

Pela instrumentalização, podemos sentir mais coisas, e, se podemos senti-las, a nossa tendência é tentar armazená-las. (Eaton, Deutsch & Deroos, 2012). Este é o desafio que está muito presente nos nossos dias, as organizações estão diante de um grande volume de informação, mas não sabem como extrair valor dela, pois a maior parte está em formato não estruturado, de maneira que é difícil saber o que vale a pena guardar. Por conta dessas mudanças, surgiu o conceito de “*Big Data*”.

Pretende-se investigar o conceito de “*Big Data*” e explorar o seu potencial como ferramenta capaz de gerar valor aos serviços das organizações nos seus mais diversos segmentos. Usando as ferramentas ideais e realizando um bom trabalho de seleção de dados é possível extrair informações valiosas que levam a uma série de melhorias e oportunidades. Com a necessidade de converter todo e/ou qualquer dado em informação faz com que as técnicas sejam aprimoradas e ao mesmo tempo colocadas em prática a fim de alcançar os objetivos desejados.

Através deste trabalho apresentarei as características de *Big Data*, as suas vantagens, oportunidades e desafios, qual a sua influência e quais as mudanças que serão necessárias nos sistemas de análise para compreender o que está por vir. Torna-se necessário pensar em formas de analisar e processar esses dados. (Davenport, Bart & Bean, 2012).

É importante tomar nota de que os dados apresentam características diversas, estes podem ser estruturados ou não estruturados. Esta diferença de grupos de dados é respeitante

à integração de dados provenientes de diferentes fontes de dados, para extrair, validar e transformar os dados de forma a auxiliar o processo de tomada de decisão. (Russom, 2013).

Como qualquer estudo, também este deve ter como ponto de partida uma questão que é crucial para o volume de dados armazenados, os quais estão a crescer a olhos vistos e em relação aos quais as técnicas tradicionais de exploração de dados não são mais adequadas para tratar a grande maioria: Quais as características de um ambiente de dados para recolher, armazenar e disponibilizar dados para posteriores análises? Assim, para responder a essa indagação, o objetivo desse estudo é compreender, através da visão de auditores, como se dá a relação das dimensões de *Big Data* no processo da auditoria.

Neste contexto, o termo “auditoria de dados” não é ainda amplamente utilizado na indústria das Tecnologias de Informação para Auditoria (TIC). Talvez por isso, esta prática esteja ausente em muitos processos de negócio em organizações de diversos segmentos. Assim, é um interessante ponto a ser investigado, principalmente levando-se em conta o atual momento em que aplicações tais como as de inteligência de dados e exploração de ambientes de *Big Data* estão na ordem do dia. A auditoria dos dados deve ser tratada numa perspetiva diferente de um projeto tradicional e deve, sim, ser entendida como um processo pontual, para resolver um problema específico de negócio, ou até mesmo para apoiar uma determinada fase de um projeto tradicional.

Big Data abre-nos as portas da EI -“Intelligent Economy” ou economia inteligente com base no fluxo contínuo de informações. (Taurion, 2012). Quer-se com isto dizer que este conceito veio revolucionar os termos, que tão bem conhecemos, competitividade e competência, são pontos essenciais no mundo das organizações.

1.1 Motivação e pertinência

A motivação para a realização deste trabalho centrou-se em conhecer de forma mais profunda o cenário da utilização de dados, em particular, no contexto *Big Data* e perceber o impacto que tal poderá ter na forma como os auditores realizam as suas tarefas. Desta forma, surgem desafios em analisar informação de forma rápida e eficaz, de maneira a trazer vantagem para quem a analisa. Neste caso de estudo, e como o tema de dissertação “Auditoria em tempos de *Big Data*: Requisitos mínimos de controlo” procura perceber-se é possível efetuar análises de grandes volumes de dados provenientes de diversas origens e diferentes estruturas de dados, procurando viabilizar e validar toda a análise de informação em ambiente de auditoria. A crescente importância e desafios da utilização das tecnologias de informação no dia-a-dia do auditor justificam a relevância do tema.

Num mundo onde voga o conceito de *Big Data* há imensas perguntas por responder e padrões por descobrir, tornando-se uma chave de pesquisa saber como responder a essas questões da forma mais correta.

1.2 Objetivos do trabalho

O objetivo principal e concreto do presente documento visa estudar um contexto de *Big Data* em Auditoria. Assim sendo, pretende-se com a dissertação um cruzamento de informação entre o Controlo Interno (CI) e *Big Data* num âmbito real, procurando entender os desafios que representa na análise de informação. O objetivo geral é caracterizar e exemplificar o termo *Big Data* com o intuito de compreender o seu surgimento e aplicação no atual cenário de dados gerados de forma contínua e aleatória.

A dissertação deverá permitir como resultado uma reflexão de um estudo de caso, quais os objetivos de controlo mais relevantes a implementar / avaliar para responder às atuais necessidades do mercado.

Em linhas mais abrangentes também terá como objetivo oferecer uma introdução ao assunto *Big Data & Analytics*, identificar os principais desafios para sua efetiva utilização.

Descrever em linguagem simples, inteligível a todos, o que realmente é *Big Data*, procurando desmistificar que os meios de comunicação têm noticiado como a solução para a inteligência do mercado.

1.3 Principais contributos deste trabalho

O *Big Data* é uma das grandes tendências apontadas a curto e médio prazo nas organizações. Há cada vez mais a preocupação de apresentar soluções para dar resposta a esta tendência, à análise metódica dos dados para tomar melhores decisões. A quantidade de dados torna-se menos relevante quando existe eficiência na *Analytics*.

Muitas organizações ainda estão céticas quanto à adoção de soluções de *Big Data*, identificando alguns obstáculos a ultrapassar, nomeadamente em matéria de qualidade dos dados, de segurança e de formação dos utilizadores. O tema é pertinente no sentido de sensibilizar para avanços na tecnologia para gerir o *Big Data* em diversas áreas, sendo que, neste caso em concreto, se estuda a aplicabilidade no contexto da auditoria. Quanto maior a utilização de soluções *Big Data* maior a confiança na análise dos dados como base para a tomada de decisões.

1.4 Estrutura do relatório

O presente texto introduz o tema *Big data & Analytics* e Auditoria e encontra-se dividido em cinco capítulos respeitantes ao assunto:

O primeiro capítulo corresponde à informação introdutória e contextualização dos restantes capítulos.

O segundo capítulo refere-se à revisão da literatura “Estado da Arte”, introduzindo os conceitos essenciais associados ao *Big Data* e à *Analytics*, e assegurando a “ponte” entre a realidade do processamento de dados e o fenómeno *Big Data* bem como o conceito no

contexto da Auditoria. Ainda neste capítulo são clarificados os conceitos, partindo das definições e perspectivas de diversos especialistas na área.

No terceiro capítulo é mencionado a metodologia que foi empregue na dissertação, neste caso é referido o método – “Estudo de Caso” com a respetiva estratégia de investigação.

Por fim, capítulo cinco, estão expostas as conclusões desta pesquisa, Apresenta-se também uma reflexão sobre temas para trabalhos futuros que podem vir a ser desenvolvidos por investigadores interessados nesta temática e as limitações que se fizeram sentir ao longo desta dissertação.

2 Estado da Arte

2.1 *Big Data*

Com o expansivo crescimento de *Big Data* e *Analytics*, as organizações estão a procurar formas de proteger e transformar grandes volumes de dados brutos em *insights* inteligentes. (Taylor, 2011). Torna-se fundamental resolver os problemas de segurança em ambientes de *Big Data* logo desde o início. Onde há dados, há o potencial de violação da privacidade dos dados, acesso não autorizado ou acesso inadequado por parte de utilizadores privilegiados. De facto, conforme os ambientes de *Big Data* ingerem mais dados, as organizações enfrentam riscos significativos e ameaças aos arquivos que contêm os dados. (Preez, 2012). Na Figura 1 está exposto o excesso de informação produzida diariamente a nível mundial, que tornou-se extremamente difícil de ser analisado por, na grande maioria dos casos, os dados serem de tipo não estruturado – os quais representam 85% de todos os dados disponíveis - pelo que emerge o surgimento de novas soluções tecnológicas.

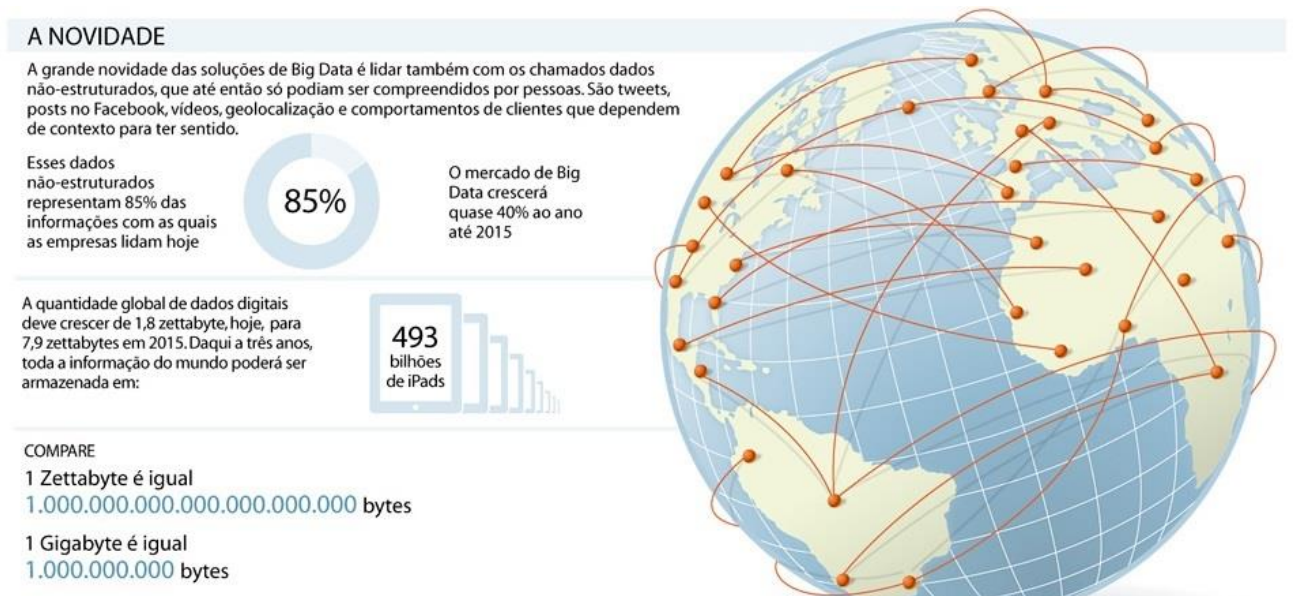


Figura 1: Dados Não Estruturados - Uma possibilidade de desenvolvimento social.

Fonte: <https://cadernodia.wordpress.com/2013/06/18/big-data-uma-possibilidade-para-o-desenvolvimento-social/>

Embora esta seja uma questão relativamente nova, tem preocupado muito as organizações. No fundo, tudo se relaciona com a quantidade incrível de informação que existe hoje. As organizações estão inundadas de dados tanto sob a forma de dados estruturados (ou seja, que estão no seu sistema de gestão) como não estruturados (filmes, comentários, áreas específicas das mensagens de email, por exemplo). (Vijayan, 2012).

O prefixo “zetta” já faz parte do léxico corrente dos sistemas de armazenamento de informação, estamos a falar de zettabytes, (Turner, Gantz, Reinsel, & Minton, 2014), ou seja, 10^{21} bytes. Um número incrível que corresponde a milhares de milhões de milhões de milhões de bytes. Ou seja, milhões de milhões de gigabytes. Foi esta enorme acumulação de informação que levou à criação da expressão *Big Data*. E, mais importante, levou ao aparecimento de um dos maiores desafios de sempre para os auditores e consequentemente para os empresários: extrair valor estratégico de toda essa imensidão de dados. A respeito disto, o autor coloca a questão como podem então os auditores e as organizações tomarem decisões (Gouveia & Ranito, 2004) rápidas baseadas em dados específicos e concretos? A expressão “*Big Data*” não trata apenas de designar grandes volumes de dados. Quando uma organização implementa uma solução *Big Data* significa que adotou mecanismos para lidar, analisar e retirar mais-valias do facto de haver tanta informação. O tratamento desta informação exige a implementação de circuitos, a criação de interfaces, a monitorização de processos. Ou, de um modo mais simples, exige preparação, trabalho e estratégia. Tradicionalmente para alguns estudiosos a tecnologia existe para assistir precisamente nessas áreas. É muitas vezes no sistema de informação que está a diferença entre analisar correta ou incorretamente a informação. (Manyika, Chui, Brown, Bughin, Dobbs, Roxburgh & Byers, 2011).

É possível visualizar na Figura 2 as principais componentes do *Big Data*, preocupações e necessidades do *Big Data*. Todos os elementos são importantes e devem ser considerados num projeto de *Big Data*.

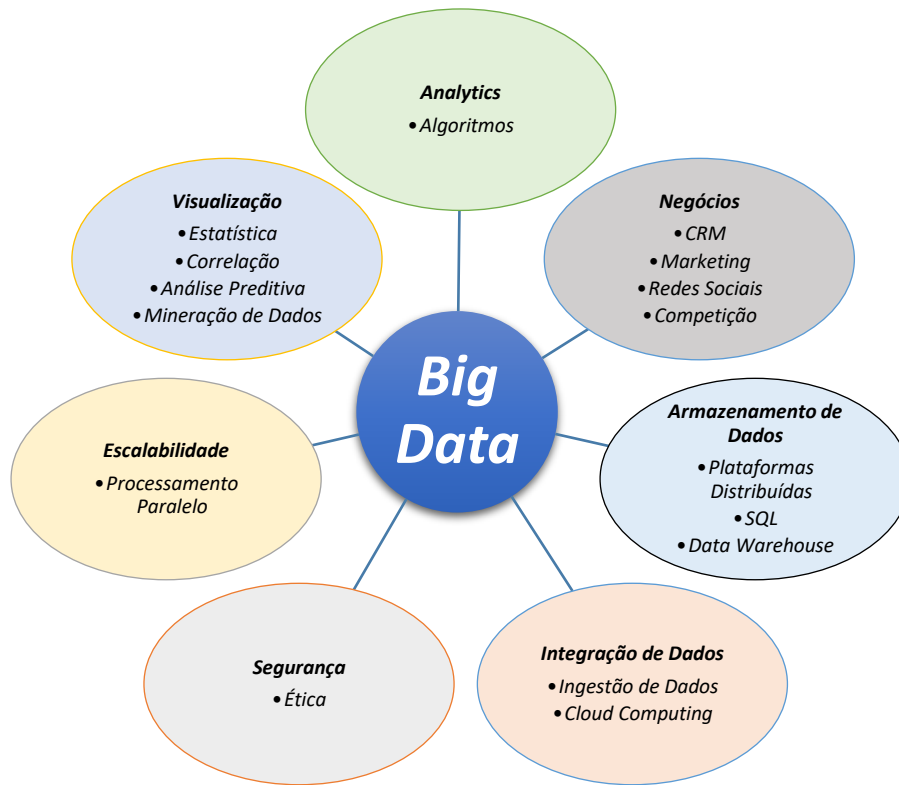


Figura 2: Visão Geral *Big Data*.

Fonte: <http://corporate.canaltech.com.br/coluna/big-data/Visao-geral-da-Arquitetura-de-Big-Data/> Adaptado.

Quanto mais dados surgem, potencialmente mais informação e conhecimento se gera, mas mais difíceis se torna aceder e processar os dados. E, principalmente, de os usar para tomar as melhores decisões. É por isso que as organizações têm mesmo de estar preparadas para esta nova realidade. As que melhor trabalharem esta informação, ou seja, que produzirem mais rapidamente resumos e sínteses para tomar decisões, vão adiantar-se em relação às outras. Muitas vezes o que acontece é que a informação até já existe nos sistemas, mas como os humanos a desconhecem não puderam decidir sobre ela e perderam assim a oportunidade. Apenas os sistemas integrados têm a capacidade de identificar toda a informação útil e mostrá-la aos utilizadores. Na Figura 3 está definido o ambiente de *Big Data*, onde se pretende demonstrar que o grande foco da questão está numa boa visão e definição de negócio, para ter uma solução - ou mesmo um projeto - de *Big Data*. Saber a questão adequada para um problema de *Big Data* e vincular esta questão com uma

necessidade real da organização é fundamental para este tipo de projeto. *Big Data* pode ser uma boa ferramenta para atender este objetivo, desde que se tenha a noção da precisão onde se quer chegar.

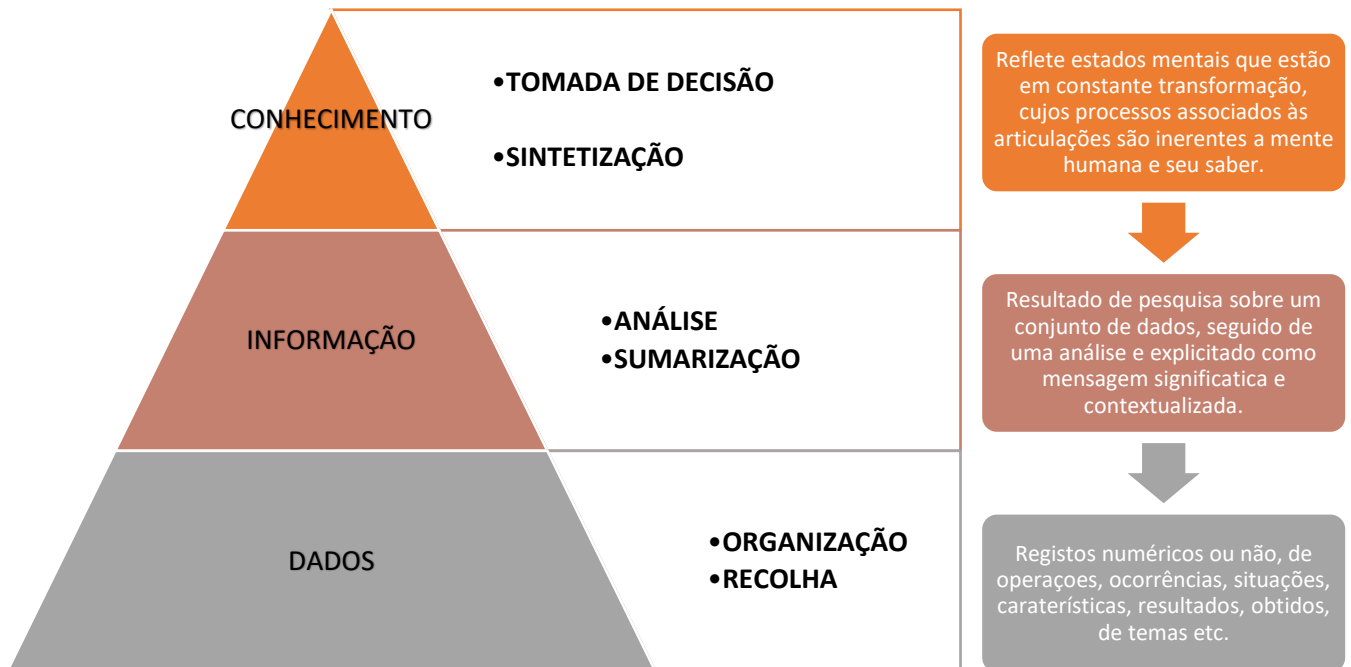


Figura 3: Dos dados ao conhecimento. Editado pela autora.

Fonte: Tipologia de Davenport (1998). Adaptado.

As organizações, públicas ou privadas, irão aumentar o volume de informação consumida, passando a monitorizar interna e externamente as suas operações, levando a um melhor fornecimento de serviços ao cliente, à melhoria das análises e à tomada de decisão.

O termo “auditoria de dados” não é amplamente utilizado na indústria de tecnologia. Talvez por isso, esta prática esteja ausente em muitos processos de negócio em organizações de diversos segmentos. Este é um interessante ponto a ser investigado, principalmente levando-se em conta o atual momento em que aplicações, tais como as de inteligência de dados e exploração de ambientes de *Big Data* estão na ordem do dia. A auditoria dos dados deve ser tratada numa perspetiva diferente da de um projeto tradicional e deve ser entendida como um processo pontual, para resolver um problema específico de negócio, ou até mesmo

para apoiar uma determinada fase de um projeto tradicional. Neste contexto, a auditoria dos dados é um “processo de apoio” fundamental para o sucesso dos diversos projetos da Auditoria. No entanto, outro autor busca a importância da auditoria de dados pode e deve fazer-se em vários processos pontuais de negócio, principalmente quando os resultados podem representar uma expressiva redução de custo e risco para a operação. (Dumbill, 2012).

É importante refletir que, desde há muitos anos, os auditores de Tecnologias de Informação (TI) têm sido capazes de confiar em ferramentas de *Analytics* relativamente elementares para realizar análises para tirar conclusões. Grande parte do trabalho envolvido na realização de auditorias implica a inspeção de dados gerados a partir de sistemas, dispositivos e outras aplicações. Estes dados incluem a configuração transacional e dados brutos de sistemas ou aplicativos que podem ser acedidos e validados, selecionados, formatados e testados de acordo com critérios pré-definidos. A utilização da tecnologia para o armazenamento de informação podem permitir aos auditores tirar conclusões-chave a partir de uma grande variedade e grande variedade de fontes de dados (internos e externos à organização). Estas conclusões, o autor assinala que podem refletir mudanças no perfil de risco global, configurarem-se em novos fatores de risco internos específicos da organização e, como distorção relevante para as demonstrações contabilísticas, o risco de fraude e risco de segurança. (Setty & Bakhshi, 2013).

2.1.1 A realidade do processamento de informação

Tão importante quanto gerar informação é a capacidade de processamento de grandes volumes de dados em alta velocidade. Um estudo realizado pela IDC indica que em 2020, o volume de informação digital em circulação, a nível mundial, deverá ultrapassar os 40.000 exabytes. Isto equivale a uma multiplicação por 10 dos valores atuais. (Turner, Gantz, Reinsel & Minton, 2014). Inevitavelmente, o desafio de analisar e entender grandes quantidades de dados está a crescer, criando novos desafios ao armazenamento. O termo *Big Data*, de forma singela, significa grandes dados, porém, esta expressão encerra desafios significativos para as organizações, sendo capaz de definir o futuro de organizações e organizações no que diz respeito à análise e estruturação de dados.

Andreas Weigend, um dos especialistas em *Big Data* e mentor do projeto Social Data Lab, refere que “antes acreditávamos que era só começar com os dados, para se chegar aos insights, que se traduziriam em conhecimento e eventualmente sabedoria. Hoje sabemos que sabedoria não é suficiente. É preciso começar com as perguntas e então chegar à decisão”. (Weigend, 2014). Mesmo pequenas quantidades de dados podem ser difíceis de gerir, tanto em termos de armazenamento como de análise. De um modo mais geral, e um ponto interessante citado por este autor, o ser humano, enquanto indivíduo, consome informação para tomar decisões e, desta forma concretizar em ação as suas intenções. A melhor prática é a verdadeira racionalização dos dados. (Gouveia & Ranito, 2004). Na Figura 4 está exposta a forma como os humanos reagem à informação e tomam as suas decisões destacando-se a presença de áreas de um Data Warehouse desenvolvido para extrair os dados adequados à tomada de decisão a partir das Bases de Dados Operacionais. A forma como o conhecimento para a tomada de decisão é construído depende da pessoa que pretende construir esse conhecimento pelo que se destacam as entidades Clientes, Gerentes e Funcionários.

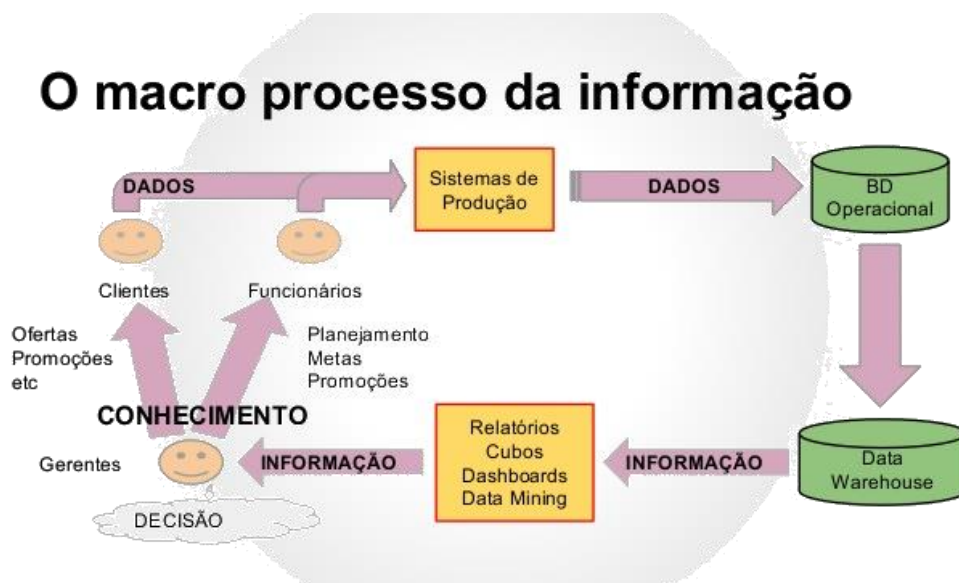


Figura 4: O Processo da Informação no contexto da tomada de decisão.

Fonte: <http://pt.slideshare.net/petroniocandido/big-data-e-data-science-admirvel-mundo-novo-iv-sic-ifnmg>

A este respeito cita o autor que uma possível definição de *Big Data* é referida como o conjunto de dados, cujo tamanho vai para além da capacidade das ferramentas típicas de bases de dados no que respeita à captura, armazenamento, gestão e análise dos dados. (Bughin, Chui, & Manyika, 2010). Define-se este conceito, também, de acordo com a

habilidade das pessoas em recolher e analisar o vasto volume de dados que estamos a gerar no mundo. (Marr, 2013).

2.1.2 O que é o Big Data

O termo *Big Data* refere-se a um grande conjunto de dados armazenados. Diz-se que o *Big Data* se baseia em 5 Vs: velocidade, volume, variedade, veracidade e valor.

O conceito *Big Data* pode ser resumidamente definido como uma coleção de bases de dados tão complexa e volumosa que se torna muito difícil (ou impossível) realizar algumas operações simples (remoção, ordenação, sumarização) de forma eficiente. Como já referido, o termo é recente embora, "*a noção mais ampla que o fundamenta não é particularmente nova*" (Riffat, 2014). Atualmente é um conceito que se poderá aplicar em praticamente todos os campos do conhecimento humano" (Schonberger & Cukier, 2013). Desta forma, é necessário conhecer algumas definições (sintetizadas em Tabela 1) estabelecidas para o conceito de *Big Data*. *Big Data* refere-se às grandes base de dados, cujo tamanho está além da habilidade de ferramentas de *software* comuns de captura, armazenamento, gestão e análise. (Manyika et al, 2011).

O termo *Big Data* é cada vez mais utilizado, sobretudo porque as organizações estão a perceber o valor e as mais-valias que podem obter através das enormes quantidades de dados que possuem, já que esses dados e as análises a realizar podem ser catalisadores do processo de inovação da organização através da adequação da realidade atual face à estratégia da organização. (Stonebraker, 2012).

Todas as organizações, umas mais rapidamente que outras, estão a perceber a importância do *Big Data*, compreendendo as mudanças necessárias a efetuar para beneficiar das vantagens competitivas da utilização do *Big Data*. (Trelles, Prins, Snir, Jansen, 2011). O crescimento, proliferação e influência das redes sociais na sociedade foram os principais fatores de influência para o crescimento e importância do termo *Big Data*. (Manyika et al., 2011).

O termo *Big Data* tem sido empregado com distintos significados, dependendo do contexto: ora se refere à prática de salvar e armazenar indefinidamente dados dos mais variados tipos, ou às próprias bases de dados com características específicas de volume e

diferentes formatos, que excedem a capacidade de processamento de ferramentas tradicionais. Ora diz respeito ao próprio fenómeno de explosão da geração de dados, pode fazer referencia também a aspetos tecnológicos e de infraestrutura e, ainda, a determinados tipos de aplicação que se beneficiam de um ambiente em que convergem diferentes tecnologias.

Apesar de ainda não haver um conceito consolidado sobre o *Big Data*, algumas das suas características aparecem mencionadas em grande parte da bibliografia referente ao assunto. O *Big Data*, mais que a convergência de um conjunto de tecnologias, oferece tanto às organizações como aos cidadãos, ou consumidores finais, um ambiente fértil à inovação, que certamente impulsionará mudanças. (Meadows, 2014).

Big Data tem uma importância inimaginável nos dias que correm:

1. Facilita ao ser humano o pensamento criativo, a perceção de *insights*, a formulação de perguntas e hipóteses;
2. Permite a cooperação entre homens e computadores, para análise, o entendimento, a tomada de decisão e o controlo de situações de alta complexidade, sem a dependência de programas e algoritmos rígidos, previamente desenvolvidos;
3. Permite a interação e troca de informações em tempo real, de forma tão ou mais eficaz do que entre seres humanos.

A capacidade para tratar *Big Data* não pode ser conseguida com uma única tecnologia, mas sim como uma combinação de varias tecnologias, umas mais antigas e outras mais recentes e que procuram ajudar as organizações a obter uma maior eficácia na gestão de grandes quantidade de dados, bem como a resolução dos problemas associados ao seu armazenamento.

Para uma visão geral sobre o conceito do *Big Data*, extrai-se da literatura definições distintas na visão de alguns autores. As mesmas são apresentadas na Tabela 1.

Tabela 1: Definições para o termo *Big Data*.

<i>Big Data é...</i>	<i>Autor</i>
<i>"(...) Big Data= Volume + Variedade + Velocidade + Veracidade + Valor"</i>	TAURION (2012)
<i>"As tecnologias de Big Data descrevem uma nova geração de tecnologias e arquiteturas projetadas para extrair economicamente o valor de volumes muito grandes e de uma grande variedade de dados, permitindo alta velocidade de captura, descoberta e/ou análise".</i>	GANTZ & REINSEL (2011)
<i>"Big Data refere-se a conjuntos de dados cujo tamanho vai além da capacidade das ferramentas de software de bases de dados típica para capturar, armazenar, gerir e analisar."</i>	MANYIKA, CHUI, BROWN, BUGHIN, DOBBS, ROXBURGH, AND. BYERS (2011)
<i>"Big Data, em geral, é definido como ativos de alto volume, velocidade e variedade de informação que exigem custo-benefício, formas inovadoras de processamento de informação para maior visibilidade e tomada de decisão."</i>	MOHAPATRA, PARISA, AND BANERJEE (2013)
<i>"Big Data é um termo genérico para dados que não podem ser contidos nos repositórios comuns; refere-se a dados volumosos demais para caber num servidor; não estruturados demais para se adequarem a bases de dados organizadas em linhas e colunas; ou fluídos demais para serem armazenados em data warehouses estáticas."</i>	DAVENPORT (2012)

Fonte: Elaboração Própria

Perante as diversas definições de *Big Data* encontradas, neste trabalho optou-se por utilizar a seguinte por considerar a mais esclarecedora para os propósitos deste trabalho: *Big Data* é mais do que simplesmente uma questão de tamanho, é uma oportunidade de encontrar *insights* em novos e emergentes tipos de dados e conteúdos, para tornar o negócio mais ágil e para responder a perguntas que foram anteriormente consideradas fora de seu alcance.

Seguindo a fundamentação teórica deste autor, e observando que foram apresentadas citações de autores mais recentes, que vêm estudando o tema nos últimos anos. Assim, com base na tabela acima, pode-se notar uma similaridade nos conceitos, os quais podem ser sintetizados como sendo um grande volume de dados estruturados ou não, de fontes diversas, que devem ser geridos e analisados de forma particular. As organizações estão, atualmente, perante cinco desafios. (Birchield, 2013):

1. Liderança: As equipas de liderança devem definir objetivos claros, e fazer as perguntas certas. Grande performance ao nível dos dados não apaga a necessidade de visão e/ou visão humana;
2. Gestão de Talentos: Deve-se atrair profissionais qualificados para trabalhar com grandes quantidades de informação (Waller, & Fawcett, 2013);
3. Tomada de Decisões: As organizações têm acesso a informações para garantir o processo de decisão de forma eficaz;
4. Tecnologia: Deve-se usar as ferramentas necessárias para lidar com o volume, velocidade e variedade de *Big Data*.
5. A cultura da organização: As organizações de maior dimensão perguntam “o que sabem” e não “o que eles pensam”.

Neste sentido, baseado no estudo dos autores, para gerir e extrair informações significativas a partir da grande quantidade de dados atual existe uma necessidade crescente de se pensar em técnicas e sistemas de *Analytics*. (Esteves, Hacker & Rong, 2014). Estes autores nos remetem a pensar que muitas organizações estão ainda indecisas em adotar sistema de análise de *Big Data*. (Chen, Chiang, & Storey, 2012).

A informação armazenada na base de dados é conhecida como dados estruturados, porque é representada num formato estrito. Por exemplo, cada registo numa tabela de base de dados relacional. Quando se lida com um enorme volume de dados, o mais comum é que provenham de diferentes fontes e se apresentem com formatos distintos entre si. É um grande desafio extrair, limpar, organizar, correlacionar, ligar e transformar esses dados em informação relevante. Para que tal seja possível, é necessário criar uma base de dados onde constem todos os dados de uma base de dados, capturado através de aplicações diversas, criadas com a finalidade de orientar estes dados, armazenando-os de forma clara e organizada, facilitando, assim, a busca e manipulação desses dados. Na Internet, apenas 25% dos dados estão armazenados em bases de dados.

Da pequena parte que se concentra em bases de dados, a maioria destas informações pertencem a organizações que precisam de um pouco mais de segurança e organização. *Big Data* afeta diretamente ou indiretamente a vida de qualquer pessoa até mesmo em situações comuns, como ir para o trabalho e na vida pessoal. Definitivamente o *Big Data* é fundamental em tempos como hoje. Sem ele, num curto espaço de tempo, a informação

tenderia a ser perdida e/ou mal utilizada. Assim, pode dizer-se que o *Big Data* ajuda as organizações a serem mais eficientes já que podem usar dados para tomarem decisões de sobre a informação analítica que permitem que operem melhor e mais rapidamente. É preciso ter uma estratégia para explorar dados no contexto de cada organização, e, frequentemente, é necessário relacionar dados de clientes, de fornecedores estruturados ou não estruturados. Assim, graças a essas análises é possível realizar análises preditivas, por exemplo.

No seguimento da Figura 5, os dados não estruturados podem encontrar-se em documentos, arquivos, gráficos, imagens, textos, relatórios, formulários ou gravações de vídeo ou áudio que não tenham sido codificados, ou de outra forma estruturados, em linhas e colunas ou registos. De acordo com muitas estimativas, cerca de 90% de todos os dados armazenados são não estruturados. Dados não estruturados podem ser também descritos como dados que se encontram armazenados num determinado contexto. (Claro, 2012).

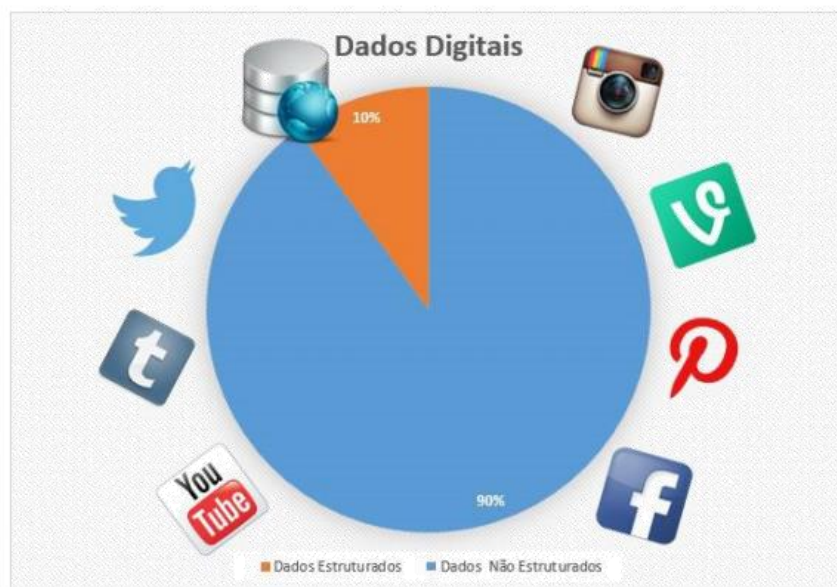


Figura 5: Dados Estruturados VS Dados Não Estruturados.

Fonte TAURION, 2014.

2.1.3 Atributos do *Big Data*

O processo de gerar, armazenar, processar, analisar e utilizar os dados é categorizado por especialistas da área em cinco partes denominadas “cinco V’s”. No entanto, nem sempre todos os V’s estão presentes nos casos de uso de *Big Data*, e, mesmo estando presentes, podem não se apresentar sempre com as mesmas características.

Conforme o artigo dos autores, *Big Data* também proporcionou um crescimento na complexidade dos dados, fazendo com que os atuais sistemas de gestão de base de dados tenham dificuldade em os armazenar. (Manyika et al, 2011). O conceito de *Big Data* pode ter várias interpretações, mas é importante realçar que está a assente no princípio das três dimensões iniciais: “Volume”, “Variedade” e “Velocidade”, às quais se juntaram mais duas dimensões “Veracidade” e “Valor” formando, assim, a designação 5 Vs, sistematizando, desta forma, parte do trabalho realizado pela comunidade científica e académica que investiga nesta área. (Stonebraker, 2012).

Sistematizam-se agora os atributos que têm sido descritos na literatura como sendo os mais comuns associados a *Big Data*:

- **Variedade** – Os dados são gerados em todos os tipos de formatos - de dados estruturados, dados numéricos em bases de dados tradicionais, até documentos de texto não estruturados, e-mail, vídeo, áudio, dados de cotações da bolsa e transações financeiras. Ainda podemos classificar os dados em semi-estruturados como, por exemplo, os ficheiros XML. (Schaeffer, & Olson, 2014).

Variedade é a forma em que os dados se encontram, divididos entre estruturados e não estruturados, sendo, na sua grande maioria não estruturados em diversos formatos. Para os dados não estruturados há dificuldades de análise (para extração de *insights*) e também de armazenamento (uma vez que dados não estruturados geralmente ocupam mais espaço digital). (Davenport, 2012). Corresponde à utilização de diferentes estruturas de dados: estruturadas, não estruturadas e semiestruturadas.

A dimensão variedade também diz respeito às características do tipo de dados no que concerne à sua fonte/proveniência. (Davenport, 2012). Um dos desafios das organizações é ter a capacidade de extrair os dados das diversas fontes existentes, de forma contínua e por padrões, para que a informação seja oportuna. Porém, esta análise de *Big Data* tende apenas

a demonstrar o que os utilizadores fazem, mas não a razão pela qual procederam de determinada forma. Outro aspeto importante a ser ressaltado na *Analytics* em *Big Data* é o desenvolvimento de *standards* e procedimentos para o uso de dados, tentando, dessa forma, gerir menor diversidade de fontes. (Zimmer, 2010).

- **Volume** – caracteriza a quantidade de dados que é gerada continuamente (na ordem de terabyte, petabyte, exabyte, zetabyte). Um estudo realizado em 2014 pela IDC revela como o surgimento das novas tecnologias tem contribuído para o aumento do volume de dados. (Kuenkaikaew, 2013).

Diariamente, grandes quantidades de dados são geradas e processados por utilizadores e organizações, na ordem dos petabytes de dados, e estima-se que este valor dobre a cada 18 meses. Tal representa uma grande quantidade de dados a ser recolhida e analisada. (Chede, 2012).

A dimensão volume é a que aparece com maior frequência nas definições de *Big Data*, justamente pela elevada quantidade de dados gerados nos últimos anos de forma gradual. A quantidade de informação crescerá até 50 vezes entre os anos de 2010 e 2020 e totalizará um crescimento de 300 vezes se considerado os anos de 2005 a 2020. (Gantz & Reinsel, 2012). Esse grande volume é oriundo de novos equipamentos capazes de gerar grandes quantidades de dados. Caminhamos, assim, para os zettabytes.

- **Velocidade** – caracteriza a velocidade a que esses dados são processados (nomeadamente em tempo real e *streams*), ou seja, a rapidez na troca de dados e troca de informação embora esteja também relacionada. (Mohapatra, Parisa & Banerjee, 2014).

Com a rapidez com que os dados devem ser analisados, sendo que, em muitos casos, tal poderá corresponder a análises em tempo real. Sendo que o volume de informação que é gerado de forma rápida e crescente, o que diminui o espaço de tempo entre as tomadas de decisão, o grande desafio aqui é conseguir recolher e armazenar os grandes volumes de dados em tempo útil, procurando utilizar os dados históricos e em tempo real para suportar as decisões operacionais.

A velocidade dos dados diz também respeito à velocidade de criação dos dados. Uma análise imediata dos dados obtidos pela organização permite que esta seja mais ágil do que

os seus concorrentes. (Mcafee & Brynjolfsson, 2012). A identificação dos dados deve ser feita em tempo útil, muitas vezes em tempo real. Se o tamanho da base de dados for um fator limite para a velocidade de entregados dados, o negócio pode ser prejudicado, fazendo com que o sistema possa vir a ser inundado pelo fluxo de dados e, pior, que os dados possam estar desatualizados antes de serem processados.

• **Veracidade** – a qualidade dos dados e informação é característica essencial para que os utilizadores interessados (na gestão executiva, na gestão pública e na sociedade em geral) usem e (re)usem os dados de forma apropriada e real, gerando informação e, consequentemente, conhecimento. As organizações necessitam de desenvolver melhorias na área de gestão de dados (*Data Governance*) para garantir a veracidade dos dados provenientes das novas fontes. (Mattsson, 2014).

Veracidade diz respeito à autenticidade, origem, reputação e disponibilidade das informações, sendo fundamental que estas estejam corretas e que façam sentido. (Chede, 2012). Muitas vezes os dados serão tratados e correlacionados antes de serem apresentados, e estes algoritmos devem ser bem testados e oferecer um alto nível de autenticidade para os dados, pois alguns erros de cálculo podem ocorrer.

Permite a classificação das diversas fontes de dados (estruturados, não estruturados e semiestruturados) de acordo com a sua qualidade, tomando em linha de conta aspetos como: precisão e atualidade dos dados fornecidos. É necessário que os dados sejam autênticos e façam sentido. No entanto, vários analistas argumentam que nada disso terá utilidade sem o quinto “V” – “Valor”. (Beulke, 2011).

• **Valor** – Para extrair valor a partir de *Big Data*, a informação necessita de ser verificada para determinar a sua veracidade, de forma que seja possível evitar incerteza. Acredita-se que o valor real dos dados é como um iceberg que flutua no oceano: apenas parte dele é visível no princípio, enquanto boa parte permanece oculta sob a água. As organizações inovadoras capazes de entender isso extraem o valor oculto – e por isso também, o mais interessante - e criam benefícios significativos. (Schonberger & Cukier, 2013). Corresponde à importância que os dados utilizados terão nas decisões a tomar.

Embora outros autores discordem desta ideia, o autor afirma que considera-se que para um sistema ser considerado de *Big Data* tem que possuir pelo menos qualquer conjunto composto por dois dos 5V's referidos. (Russom, 2013). Constata-se que nem todos os

sistemas de gestão de bases de dados relacionais devem ou podem ser considerados *Big Data*. (Manyika et al., 2011).

Os atributos anteriormente definidos, e refletidos na Figura 6, identificam os principais desafios que precisam de ser considerados quando se define o termo *Big Data* e a sua análise (*Big Data & Analytics*). “Valor” é a importância do resultado destes dados para as organizações: os dados passam por um filtro permitindo obter relatórios relevantes para tomada de decisões, podendo ser até um processo automática ocorrendo num sistema programado para o efeito.

Quanto maior for o nível de informação, maior será o valor de *Big Data*. Porém essa informação vai se tornando mais difícil de obter e tornando o processo cada vez mais moroso. É fundamental que os dados acrescentem valor aos futuros utilizadores, para que o enorme investimento necessário para *Big Data* não seja uma despesa inútil.

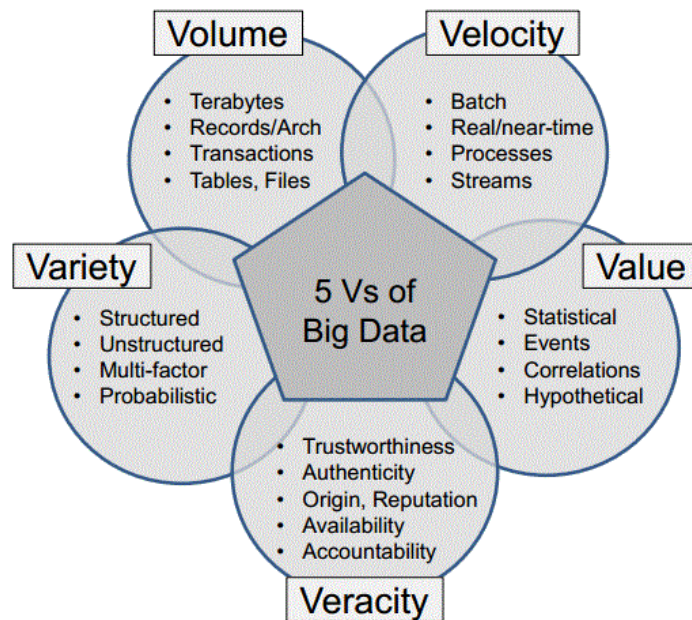


Figura 6: Os 5 V's *Big Data*.

Fonte: <http://bigdata.black/featured/what-is-big-data/>

De acordo com a fonte, acredita-se que os estudos e as pesquisas devem verificar se a presença ou não das características do *Big Data* nas organizações estão relacionadas com os processos das mesmas, entre eles o processo de decisão.

2.1.4 Por detrás de uma grande decisão está sempre uma grande pergunta

O professor Andreas Weigend, da Universidade de Stanford, afirma que *Big Data* é inútil sem boas perguntas o que nos leva a refletir sobre as questões mais relevantes e úteis que podem ser respondidas com o apoio dessa tecnologia e o nível de maturidade das organizações para elaborar e encontrar respostas a essas perguntas. A visão deste investigador passa por, no futuro, considerar que qualquer decisão tomada por uma organização poderá (e deverá) ser apoiada em dados, números e informação obtida usando técnicas de exploração de *Big Data*. (Weigend, 2014).

Os 5Vs anteriormente referidos (volume, velocidade, variedade, veracidade e valor) são itens que devem ser questionados. Conforme a Figura 7, na nova abordagem de *Big Data* esta combinação, além de todo e qualquer outro aspeto que caracteriza uma solução de *Big Data*, pode ser inviável se o resultado não trazer benefícios significativos e que compensem o investimento das organizações que passam a adotar o *Big Data* como estrutura de *Analytics* e valor de informação. (Moffitt&.Miklos, 2013). Nos seus trabalhos o autor mostra a sua preocupação com o facto de antes de armazenar os dados *Big Data* é essencial saber para que servem, como funciona, qual a capacidade de armazenamento, quais os dados da organização que devem ser armazenados e, sobretudo, ter conhecimento do serviço oferecido. Assim, se começarmos com uma boa pergunta podemos ser mais assertivos na busca dos dados e, conseqüentemente, obter um melhor resultado no projeto. (Franks, 2012).

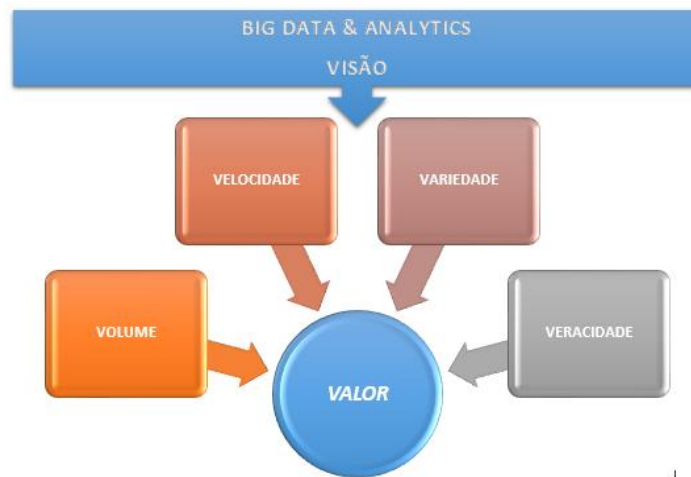


Figura 7: Nova abordagem de *Big Data*.

Elaboração Própria

2.2 *Analytics*

Encontrou-se autores que exprimem que há uma imensa variedade de necessidades nos negócios de hoje para uma *Analytics* eficaz. Por esta razão, os gestores têm que ter sensibilidade em determinados fatores que possam considerar essenciais na condução da *Analytics* e seguir um programa eficaz de análise. (Setty & Bakhshi, 2013). As tecnologias para manipulação de tal quantidade de dados são recentes e muitas organizações ainda estão em processo de aprendizagem relativamente a essas questões. É importante realçar que há um ciclo em função dos dados, isto é: os dados são gerados, armazenados, processados, analisados e, por fim usados com algum propósito. Daí ser importante ter em consideração que a fase de análise dos mesmos pode gerar novos conhecimentos para as organizações.

A *Analytics* é definida como sendo a arte ou a ciência de descobrir e analisar os padrões, identificando anomalias, e a capacidade de extrair outras informações subjacentes nos dados ou relacionados com o objeto de uma auditoria através de análise e visualização, com o propósito de planear e executar uma auditoria.

2.2.1 Âmbito de aplicação em Auditoria

Algumas recomendações com relação à formação de estratégias são passadas pelo autor, o auditor deve determinar qual a melhor abordagem para satisfazer os objetivos da *Analytics*, isto é, dar resposta a: como são os dados, onde estão disponíveis e a melhor forma para os obter tendo sempre em mente a perspetiva de negócio. Segundo o mesmo autor, a *Analytics* pode ser eficaz na organização na fase de planificação de trabalho de campo. (Singleton, 2013).

Os conceitos relacionados com *Big Data* estão intrinsecamente ligados às necessidades de busca, recuperação e recomendação numa base de dados. Na Figura 8 está o processo pelo qual se estrutura e dá o significado aos dados, consiste na transformação dos dados em conclusões úteis e credíveis, procurando tendências, diferenças e variações na informação obtida. Todos estes processos, técnicas e ferramentas usadas são baseados em certos pressupostos e com tal tem as suas limitações.

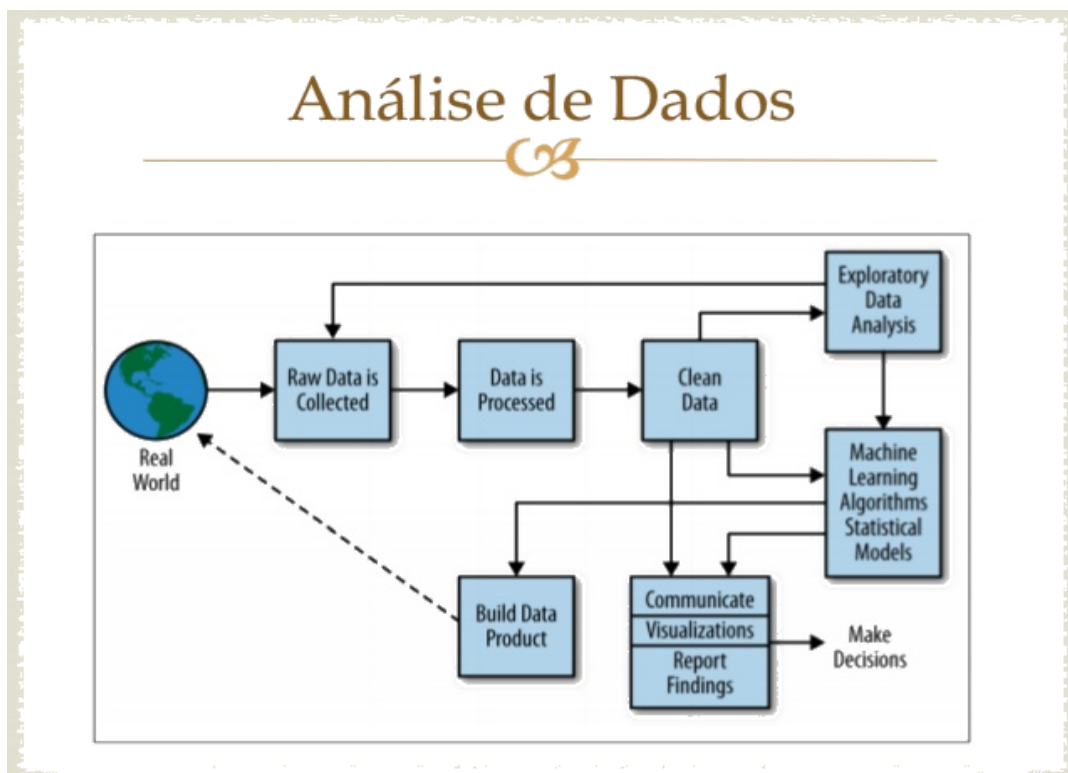


Figura 8: *Analytics*.

Fonte: <http://pt.slideshare.net/dreznik1/big-data-aplicaes-aspectos-prticos-e-cientista-de-dados>

2.2.2 A Transformação de Dados

A análise do perfil e origem dos dados é da competência da *Analytics*. É sobre esta ótica que o envio de informação para o exterior da organização é, na maioria dos casos, suportado por relatórios periódicos de operações, relatórios esses que necessitam de auditorias internas prévias ao envio da informação de modo a garantir que a informação da atividade se encontra em conformidade com os padrões esperados e com a legislação em vigor. (Singleton, 2013).

Os auditores utilizam ferramentas técnicas de *Analytics* para tirar conclusões. *Big Data* engloba não só as operações realizadas pelas organizações bem como o novo mundo de interações e observações sobre os dados. (Setty, Bakhshi, 2013). Adicionalmente, o autor coloca que a condução de auditorias implica a inspeção de dados gerados a partir dos sistemas, dispositivos e outras aplicações. (Singleton, 2013).

Muitas organizações adotam “grandes análises de dados”, procurando avanços nas técnicas de análise para a rápida expansão do conjunto de informação que as organizações têm à sua disposição para permitir uma melhor tomada de decisão. Torna-se fundamental garantir a segurança da informação. Este trabalho compete aos profissionais de risco e também aos auditores, nomeadamente aos auditores de Sistemas de Informação. Num âmbito mais profundo, este autor talvez queira revelar que o risco técnico e operacional surge como resultado da adoção da *Analytics* por parte das organizações: já a “não adoção” de *Analytics* tem influência na competitividade das organizações. (Yu, 2014). Referimo-nos a *Analytics* considerando um sistema capaz de armazenar, agregar e transformar a ampla gama de fontes de dados da organização. Neste tema, uma funcionalidade potencial é a análise de fraude, a qual representa um risco nas organizações.

Novas ferramentas e métodos de *Analytics* demonstram a possibilidade de as organizações gerarem valor a partir da existência de dados dentro das próprias organizações e de informação externa disponível. Estes avanços permitem melhor tomada de decisão e aumento da vantagem competitiva da organização/organização. No entanto, é fulcral perceber que é um investimento no *Big Data* e *Analytics* que se destina à melhoria em termos de eficácia ou eficiência de negócio. Já estes autores relatam que, esta questão envolve o binómio risco operacional versus risco de negócio. (Gantz & Reinsel, 2011).

2.2.3 Vantagem competitiva

Ao traçar um paralelo entre as ideias expostas pelos autores, a *Analytics* é fundamental para muitas organizações porque, na maioria das vezes, já foram realizados investimentos na área de análise e recolha de grandes conjuntos de dados, em que a componente analítica pode ser aplicada. (Manyika et al., 2011).

Existe uma correlação direta entre o uso de *Big Data* e a análise e rentabilidade. Existem estudos que indicam que há aumento de rentabilidade global como resultado do uso de *Big Data* de forma eficaz. A vantagem competitiva está presente caso a organização tenha as pessoas, processos e tecnologias com capacidade para fazer uso dos dados da organização. (Singleton, 2013). Para além disso, é fundamental a organização estabelecer perfis de responsabilidade, ou seja, acessos por tipo de utilizador para garantir um bom funcionamento no que à informação diz respeito. Portanto, conforme considerações feitas nos estudos do autor, algum grau de autoconhecimento organizacional é essencial, nomeadamente, pensar e apoiar os processos e identificar potenciais áreas problemáticas antes de serem realizados investimentos significativos em *Analytics*. (Mattsson, 2014).

Técnicas de *Analytics* podem conferir vantagem competitiva organizacional e para o negócio quando é tido em conta a avaliação da possibilidade de segurança ao risco de privacidade que poderia ocorrer. Os autores baseiam-se num exemplo histórico de *Analytics*: tirando partido dos dados registados, as organizações podem identificar, de forma célere, oportunidades para simplificar as operações e reduzir custos, algo que é garantido por essa visão única e sistemática dos dados. (Agrawal, Das & Abbadi, 2011).

Note-se que uma variedade de ferramentas (*softwares*) são usadas para executar *Analytics* e ajudam a identificar áreas chave de risco, nomeadamente, fraude, erros de má utilização, permitindo, assim, a melhoria dos negócios, a verificação de eficiência e eficácia dos processos e, por fim, melhoria das decisões associadas ao negócio. Sinteticamente, para investidores como estes, os benefícios para a organização incluem a eficácia, redução de erros e oportunidade de melhorar a identificação de problemas. (Turner, Gantz, Reinsel & Minton, 2014).

A *Analytics*, quando utilizada de forma eficaz numa organização, está entre os ativos mais valiosos da organização. Extraem-se benefícios se se verificarem estas três premissas:

a *Analytics* está alinhada com o modelo de negócio (*Data Governance*), o risco é gerido e o processo é planeado, projetado, implementado e testado. (Schaeffer & Olson, 2014).

No fundo o objetivo é que o conceito *Big Data* converta grandes dados em informação útil. Assim, cumprindo os princípios fundamentais no contexto de *Big Data*, as organizações podem colher os benefícios sem os potenciais riscos para a conformidade no presente e no futuro.

Pesquisas recentes nesta corrente indicam que o uso adequado de *Big Data* pode desempenhar um papel económico útil nesse sentido, promovendo inovação, competitividade e produtividade em todos os segmentos. *Big Data* pode gerar valor de diversas formas, sendo algumas delas: criar transparência entre organizações e consumidores, descobrir necessidades, expor variáveis e melhorar performance, segmentar públicos-alvo para otimizações substituir decisões humanas por algoritmos automatizados, gerar novos modelos de negócio, produtos e serviços. (Bughin, Chui & Manyika 2010).

Na Figura 9 são apresentadas as soluções para o problema do *Big Data*, cada vez mais os líderes de negócio irão recorrer ao *Big Data* como uma vantagem competitiva que lhes permite entregar serviços personalizados, tanto em pontos de encontro digitais como humanos. Nesta mesma linha de conta, o autor faz referência aos avanços em ferramentas, como bases de dados gráficas e outras técnicas emergentes de análise e classificação de informação, aportarão significado à informação desorganizada. (Gartner, 2013).

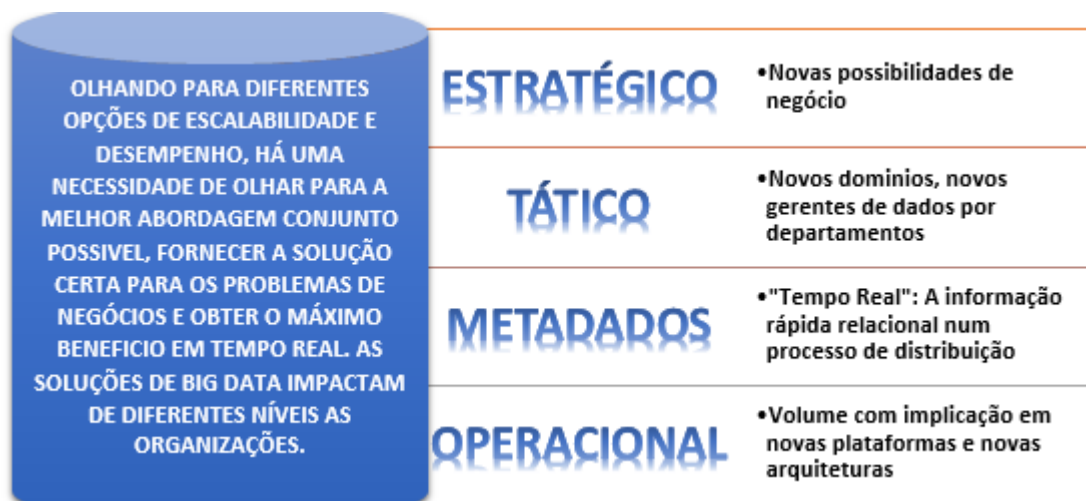


Figura 9: Soluções para o problema do *Big Data*.

Fonte: <http://maximodata.com/big-data> Editado

2.3 Auditoria - Requisitos mínimos de controlo

A quantidade de informação que hoje é produzida é muito superior à que se imaginaria no passado e o seu crescimento é exponencial. Num estudo realizado prevê-se que mais de metade das organizações portuguesas planeia a sua transformação digital.¹ (It *Insight*, 2016). Paralelamente, o estudo revela uma crescente dependência da infraestrutura tecnológica com 72% das organizações a referirem que acreditam que a procura de recursos de TI vai aumentar. O objetivo do estudo prende-se com o facto de as organizações procurarem uma melhor gestão dos dados e melhorar a capacidade de resposta às necessidades do negócio.

É de salientar que há escassez de investigadores em *Big Data & Analytics*. Há, inclusive, investigadores que sensibilizam para que se faça mais investigação porque é fundamental para sustentar futuras aplicações e ferramentas para a gestão de projetos de *Big Data* no mundo. (Boyd. & Crawford. 2012).

Voltando à auditoria, fica agora fácil perceber que o *Big Data* afetará cada vez mais a auditoria e a tornará cada vez mais dependente da análise de grandes volumes de dados produzidos pelas mais diversas fontes. Os dados passam a ser o maior ativo das auditorias.

Tendo como base este autor, a questão coloca-se: Mas de que forma é que o *Big Data* está a afetar a auditoria? (Amaral, 2014).

1. Minimização da Auditoria Tradicional: Processos de auditoria tradicional serão cada vez menos aplicados. É necessário o uso de ferramentas e técnicas de *Analytics* especializadas, pois as atuais, se assentes no modelo tradicional, serão ineficientes.² (Moffitt & Miklos, 2013).

¹ O estudo “Digital Transformation e os desafios do Datacenter”, realizado pela IDC Portugal com o apoio da Fujitsu Portugal revela que 87% das organizações portuguesas estão focadas na transformação digital. Cfr: Fujitsu dispõe de soluções que permitem abraçar o crescimento do “Universo Digital” in <http://www.itinsight.pt/news/digital/mais-de-metade-das-organizações-portuguesas-planeiam-a-sua-transformacao-digital>

² A auditoria contínua é uma inovação tecnológica do processo de auditoria tradicional que utiliza a tecnologia e automação, e embora o seu conceito já tenha quase duas décadas, na prática, ainda é uma novidade. O interesse em explorar a metodologia de auditoria contínua avançou e os profissionais estão a colaborar e a estabelecer parcerias de pesquisa com a comunidade académica, bem como, os gestores e os seus auditores reconhecem que o paradigma da auditoria tradicional está desatualizado.

2. Ineficiência das Ferramentas Tradicionais: As ferramentas tradicionais de *Analytics* não serão eficientes.³
3. Fim da Amostragem: a amostragem acompanha uma margem de erro, portanto, ele é um procedimento que só é executado pelo fato de se analisar toda a população ter um custo, que pode ser muito alto ou inviável.⁴
4. Independência também da Infraestrutura: a auditoria tem como uma de suas premissas a independência.⁵
5. Potencialização do Auditor: o *software* tradicional de auditoria define a formação padrão esperada de um auditor que analisa dados. Porém, não é tarefa do auditor selecionar e tratar dados, ele deve receber os conjuntos de dados prontos, limpos, formatados.⁶
6. Processos: a *Analytics* em si deve ter seus próprios processos e pode ser medida em termos de maturidade.⁷

Por estas razões, provavelmente, daqui a algumas décadas poderá ser falado na auditoria antes e depois do *Big Data*, já que o fenómeno poderá vir a gerar uma nova era. A auditoria de dados pode e deve se fazer presente em vários processos pontuais de negócio, principalmente quando seus resultados podem representar uma expressiva redução de custo e risco para a operação.

Todos os tipos de relatórios precisam de ser auditados internamente, antes do envio a qualquer entidade, para assegurar que as informações da operação estão em conformidade

³ Um dos importantes aspetos dessa nova realidade para a auditoria, consiste no facto dos procedimentos tradicionais de obtenção de provas de auditoria não puderem ser apropriados ou suficientes, já que a evidência em auditoria é gerada principalmente em formato eletrónico.

⁴ Inerente à amostragem em auditoria está o risco de amostragem, ou seja, o risco das conclusões do auditor, baseadas numa amostra, serem diferentes das conclusões a que chegaria se analisasse toda a população.

⁵ Afirma-se também que pode haver um auditor externo à organização, que tenha uma independência suficiente para realizar uma auditoria objetiva e real

⁶ Quando os auditores internos utilizam os componentes do método C.O.S.O. para avaliar determinado objeto de auditoria, é possível analisar se a organização possui um ambiente de controlo interno eficiente. Para isso é importante que tenha acesso a toda a informação de forma mais simplificada. Extrair e tratar dados consome no mínimo 80% do tempo de um projeto. Este cenário fica ainda mais complexo com *Big Data*: muitas fontes, com estruturas diferentes, com grande volume. O auditor deve receber seus conjuntos de dados prontos, limpo, com os fatos e dimensões relevantes do período a ser analisado, ele não deverá se envolver no penoso processo de extração e preparação. O Auditor deve focar seu esforço nos objetivos do seu trabalho, desenvolvendo a análise e a demonstração.

⁷ A natureza estratégica dessa abordagem faz com que este método se torne efetivo para alinhar os objetivos e estratégias organizacionais com os processos existentes.

com os padrões esperados e com as leis nacionais. Incumprimento de prazos e erros de não conformidade nas informações são passíveis de multas.

Ainda assim, o autor propõe que para estabelecer um processo de auditoria de dados, é importante ter em mente quatro tarefas essenciais: (Amaral, 2014)

- A primeira delas é a necessidade de se definir e documentar os processos de auditoria com as suas periodicidades, papéis e responsabilidades.⁸
- Em segundo lugar, vem a exigência de se definir os indicadores de qualidade que deverão ser acompanhados. É importante, neste caso, avaliar em valores os casos de não conformidade, estabelecendo assim o nível de assertividade e percentagem de risco que deve ser acompanhado.⁹
- Em terceiro lugar, vem a tarefa de criação dos esquemas de notificações, mediante processos de análises recorrentes, seguida da necessidade de se realizar o processo através de uma tecnologia integrável a qualquer origem de dados, que seja fácil na construção e execução dos relatórios de auditoria e que seja acessível ao utilizador.¹⁰
- Finalmente, através dos indicadores de auditoria, o fim do processo exige aprimorar de forma contínua a operação, enviando recomendações às áreas envolvidas na geração e controlo das informações.¹¹

É já conhecido por todos que as metodologias de auditoria de dados reúnem os requisitos necessários, para se tornarem num elemento obrigatório na organização, no que à exploração de dados para efeito de negócios e segurança diz respeito. Em plena era da

⁸ A este propósito a DRA N.º 300, dispõe no seu parágrafo 11: “O revisor/auditor deve desenvolver e documentar um programa de revisão/ auditoria que estabeleça a natureza, tempestividade e extensão dos procedimentos planeados necessários para implementar o plano global de revisão/auditoria. Tal programa serve como um conjunto de instruções aos membros da equipa envolvidos na revisão/auditoria e como forma de controlar e registar a adequada execução do trabalho. O programa deve também conter os objetivos de revisão/auditoria para cada uma das áreas e um orçamento de tempos com as horas estimadas para as várias áreas ou procedimentos.”

⁹ Verificar se a qualidade dos controlos utilizados efetivamente realizam sua função empírica de minimizar riscos.

¹⁰ Os papéis de trabalho da revisão/auditoria suportam a opinião do revisor/auditor a emitir no seu relatório final, mas servem complementarmente outros propósitos, como, por exemplo, contribuem para a melhoria do planeamento, facilitam o processo de verificação ou supervisão ou, ainda, servem de evidência do trabalho efetuado em sede de contestação judicial.

¹¹ Uma gestão de risco pró-ativo possibilita à organização preparar-se para o inesperado e reduzir a volatilidade negativa de indicadores contabilísticos, financeiros e operacionais.

tecnologia o auditor necessita de expressar uma opinião baseada em grandes volumes de informação, com uma estruturação complexa e em constante mudança.

É através de técnicas de *Analytics*, as quais incluem, por exemplo, a utilização de ferramentas eletrónicas, que o auditor irá conseguir analisar de forma crítica e eficiente o volume de dados que lhe é colocado à disposição por parte das organizações. (Pedrosa & Costa, 2014) Pese embora o reconhecimento dos profissionais de auditoria sobre o potencial que estas técnicas poderão ter, nomeadamente alterando a forma como a auditoria é realizada, tornando-a mais efetiva e possivelmente mais eficiente, ainda existe um longo caminho a percorrer para reduzir o ceticismo dos auditores sobre estas técnicas e assim impulsionar a sua constante utilização no processo de auditoria. As ferramentas específicas existem mas os auditores continuam a preferir trabalhar com Excel. (Pedrosa, Costa & Laureano, 2015).

Num processo de auditoria tradicional, o recurso à tecnologia ocorre em menor escala e as provas de auditoria são obtidas essencialmente em suporte físico. Neste sentido, os procedimentos substantivos assentam na sua maioria em testes de detalhes, com um maior nível de verificação física de documentação de suporte, não significa isto que com a utilização de ferramentas eletrónicas não exista qualquer tipo de análise documental, mas sim a possibilidade de reduzir o volume a analisar por via de obtenção de prova de auditoria utilizando outro tipo de procedimentos ou técnicas de amostragem. Esta abordagem é, tradicionalmente, menos eficiente (considerando projetos com uma dimensão significativa) e acarreta maiores custos para a firma de auditoria uma vez que são comuns os atrasos significativos na recolha da informação e no processamento e análise da mesma. Atualmente, este estudo justifica este tipo de auditoria que apresenta evoluções significativas, tendo as organizações de auditoria acompanhado a evolução das tecnologias.

São várias as vantagens que a utilização de técnicas de *Analytics* no decurso do processo de auditoria das quais se destacam:

1. Maior segurança nas análises efetuadas;¹²
2. Capacidade para trabalhar um grande volume de dados;¹³

¹² Identificar e analisar padrões não usuais e valores discrepantes.

¹³ Mapear e visualizar o desempenho financeiro e outros dados entre unidades operacionais, sistemas, produtos ou outras dimensões com o objetivo de focar nos riscos de auditoria;

3. Transformar a auditoria num processo contínuo;¹⁴
4. Maior eficiência na deteção de fraude;¹⁵
5. Redução de “falsos positivos”.¹⁶

2.3.1 Riscos do *Big Data*

O constante crescimento de dados pode gerar problemas quando se trata de informações pessoais: as organizações têm encontrado cada vez mais incentivos para manter uma grande quantidade de informações, para as manter durante mais tempo e para as reutilizar com maior regularidade. Cabe lembrar que a maioria dos autores são de acordo com estes que apresentam que os dados podem, aparentemente, não serem identificáveis como informação relevante mas, ao passar por determinadas análises, pode-se atingir um nível de detalhe mais minucioso da vida de uma organização. (Schönberger-Mayer, & Cukier, 2013).

O *Big Data* propicia vários benefícios, contudo a forma como os dados podem ser mal utilizados corresponde a um risco acrescido. Assim sendo, o *Big Data* pode se transformar num instrumento que pode destruir a privacidade. (Schönberger-Mayer, & Cukier, 2013). Sob este ponto de vista a preocupação com a segurança da informação é crescente, compatível com a necessidade de proteger dados confidenciais de organizações e pessoas. Embora essas questões precisem de ser levantadas, é certo que *Big Data* tem o potencial de “*transformar economias, trazendo uma nova onda de crescimento produtivo e saldo positivo para o consumidor*”. (Manyika et al., 2011).

Estes estudos caracterizam-se na ideia que os riscos e benefícios que poderão resultar do uso de aplicações com acesso às bases de dados classificadas como *Big Data* justificam plenamente o estudo.¹⁷ Na Figura 10 é apresentado a diferença na produtividade entre uma

¹⁴ Construir modelos estatísticos ou outros modelos preditivos que permitam explicar o relacionamento dos dados com outros fatores e identificar flutuações significativas no modelo

¹⁵ Identificar e analisar padrões não usuais e valores discrepantes.

¹⁶ Combinar a informação obtida de fontes diversas com o objetivo de obter conhecimentos adicionais sobre o negócio, operações ou entidade

¹⁷ Estes possíveis impactos igualmente justificam que a presente discussão seja estendida não somente aos profissionais de tecnologia e aos executivos de organizações interessados nos benefícios económicos que podem advir de vantagens competitivas oferecidas pelo *Big Data*, mas também a agências governamentais, economistas, políticos e juristas.

empresa que não utiliza *Big Data* e uma empresa que utiliza *Big Data*, é possível observar que existe uma vantagem de 5% a 6% na produtividade.

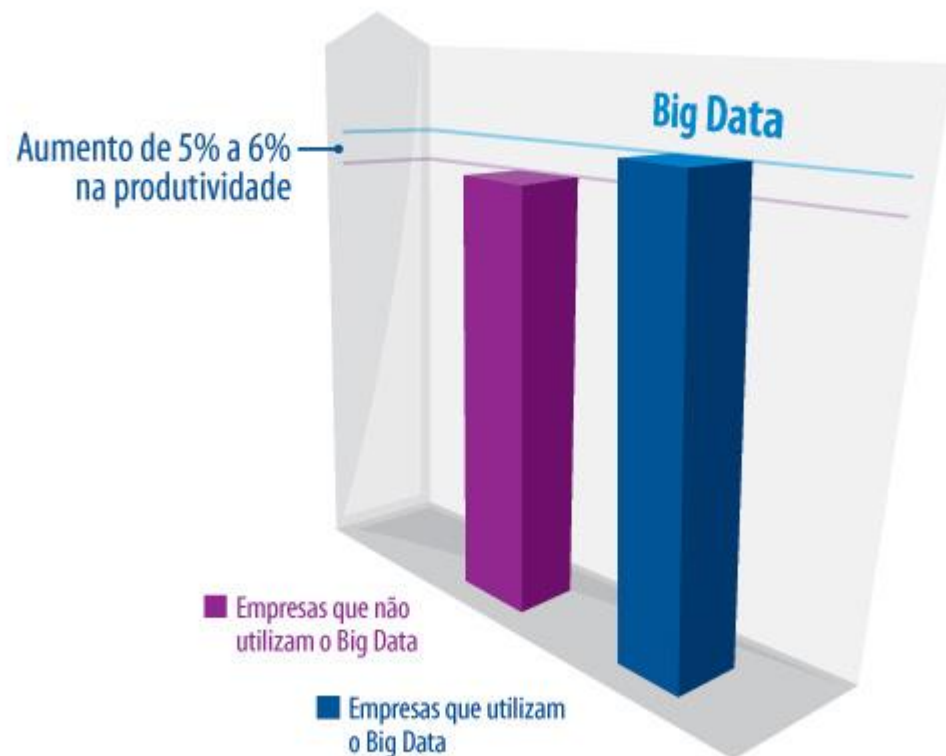


Figura 10: Riscos associados ao *Big Data*.

Fonte: <http://blogbrasil.comstor.com/bid/305729/os-riscos-do-big-data-e-a-invas-o-de-privacidade-nos-eua>

Tomando por base os conceitos tratados devemos ter em conta os possíveis impactos do *Big Data*,¹⁸ organizados como riscos, desafios e oportunidades, sob a ótica do cidadão, da indústria de TI e do mercado Empresarial em geral.

Como benefícios: (Canton, Goodman & Schneier, 2015).

1. Luta contra o crime organizado;
2. Combate a fraudes financeiras;

¹⁸ Este risco pode ser reduzido através de um adequado planeamento e supervisão do trabalho, bem como, no caso de organizações de auditoria, pela implementação de um controlo interno de qualidade.

3. Combate a epidemias;
4. Controlo preventivo de grupos de riscos de determinadas doenças
5. Localização e atendimento imediatos em casos de mal súbito ou emergências diversas
6. Planeamento de investimentos em infraestrutura
7. Planeamento de ações em caso de acidentes ou catástrofes naturais.

Entre os principais benefícios que as organizações nacionais encontram na transformação digital do negócio é a eficiência operacional, um relacionamento melhor com os clientes e a criação de novas fontes de receita.

A indústria de TI tem procurado responder adequadamente aos desafios do fenómeno do *Big Data*, lançando tecnologias específicas para atender às necessidades de infraestrutura e de técnicas para o armazenamento e a análise dados. Algumas dessas situações desafiadoras apresentam tanto componentes de risco como oportunidades de crescimentos e vantagem competitiva. Fica claro que para estes autores, a resposta da indústria de TI a esses desafios traduz-se no lançamento de diversos modelos de base de dados apropriados ao armazenamento de dados não estruturados, projetados para uma dimensão compatível com as necessidades do *Big Data*. (Guolinag, Beng, Jianhua, Jianyoung, & Lizhu, 2008). O problema desse tipo de solução é que, na prática, os relatórios analíticos são produzidos com base em dados estáticos,¹⁹ refletindo a realidade de horas ou dias anteriores.

Como desafios, riscos e oportunidades devemos ter em conta que:

1. Desafio - Desafios de ordem técnica ainda a serem solucionados de forma eficiente;
2. Risco - Pessoal qualificado insuficiente;²⁰
3. Risco - Crescimento acelerado de aplicações e do surgimento de novas organizações, agravando a competição tanto por mercados específicos como por mão-de-obra qualificada;

¹⁹ Assim a indústria de TI já oferece ferramentas voltadas para o tratamento analítico de dados, compatíveis com as características do *Big Data*.

²⁰ Pessoal despreparado para explorar as novas tecnologias. Pessoal despreparado para a análise eficiente da grande quantidade de dados.

4. Risco - Momento histórico da volatilidade tecnológica, com risco de surgimento e rápido desaparecimento de inúmeros produtos e, eventualmente, das suas respetivas organizações;
5. Benefício - Momento de disputa intensa entre as organizações bem-sucedidas com o lançamento de novos produtos, as quais com o tempo deverão ser incorporadas por outras maiores, restando, ao final desse período de maior instabilidade, apenas as organizações que melhor conseguirem se adaptar ao novo cenário tecnológico e económico.²¹

Acredito que as melhores oportunidades para obtenção de vantagens competitivas com o *Big Data* estarão centradas na capacidade de reconhecimento de padrões, principalmente na identificação de padrões adaptativos.

A IDC e a CISCO deram a conhecer os principais resultados do primeiro estudo português dedicado a avaliar a segurança no seio das organizações, que inquiriu mais de 200 médias e grandes organizações de todos os sectores: a maioria destas não sabem que estão a ser atacadas, gastando aproximadamente 12 por cento do seu orçamento de TIC em segurança, valor inferior à média europeia.²² O objetivo do estudo centrou-se em compreender os riscos, as tecnologias e as despesas associadas à área de segurança de informação, além de outras questões de gestão.²³

Big Data & Analytics representam a tecnologia de informação que apoia as organizações consciencializando daquilo que é o mercado e quais as tendências ao nível da segurança de informação. Na Figura 11 está presente o risco das organizações, seja qual for a sua área de negócio, e são determinantes na execução e alcance dos objetivos a todos os níveis, por isso a sua gestão é um imperativo. Os riscos são considerados em termos de fontes de incerteza que podem ser tratados para maximizar os ganhos, minimizando a perda.

As organizações têm de dar importância a um sistema de gestão de risco, que permita gerir, de forma eficaz e eficiente, os riscos que poderão comprometer os objetivos estratégicos da organização.

²¹ Necessidade de reagir rapidamente, atuando no mercado, preferencialmente, de forma proactiva.

²² Cfr. “Organizações portuguesas têm pouca maturidade no que respeita à segurança.” – in *itinsight.pt* estudo intitulado “Segurança de Informação nas Organizações Portuguesas 2016” foi conduzido pela IDC Portugal de forma independente e apoiado pela CISCO.

²³ “CISCO e IDC vão avaliar a segurança das organizações portuguesas” in *IT Insight*

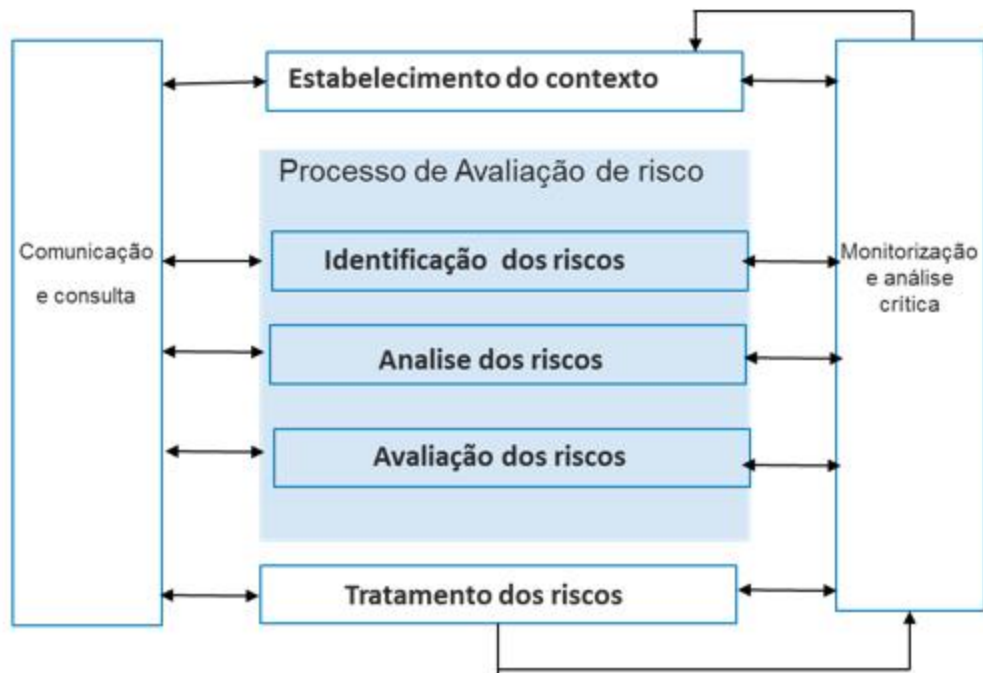


Figura 11: Análise de Riscos.

Fonte: <http://businessAnalytics.pt/iso31000/>

A "agitação" em torno do *Big Data* tem-se concentrado na infraestrutura que suporta o volume extremo, velocidade, variedade e recursos analíticos em tempo real que a infraestrutura proporciona. Onde há dados, há o potencial de violação da privacidade dos dados, acesso não autorizado ou acesso inadequado por parte de utilizadores privilegiados. (Sudhakar Sathiyamurthy, 2011). Quem é responsável pela segurança de dados na sua organização, talvez tenha que responder a perguntas como: 1) Quem faz solicitações específicas de *Big Data*? 2) Quais as tarefas de Mapear/Reduzir dados que estão a ser realizadas? 3) Existem tentativas de *download* de dados sensíveis ou é apenas uma consulta de marketing normal para obter *insight* sobre os clientes? 4) Há um número excessivo de exceções de permissão do acesso a um determinado ficheiro que tem características confidenciais, o que poderá estar a ser causado por um *hacker* numa tentativa de obtenção de acesso aos dados sensíveis por meio de algoritmos? 5) Essas tarefas fazem parte de uma lista de aplicações que estão autorizadas a aceder aos dados? 6) Foi desenvolvida alguma aplicação nova cuja existência seja desconhecida e os acessos tenham sido "herdados" de outras?

O auditor²⁴ deve executar procedimentos de avaliação do risco que proporcionem uma base para a identificação e avaliação dos riscos de distorção material ao nível das demonstrações financeiras e ao nível de asserção. Neste sentido o revisor/auditor deverá utilizar a documentação preparada no conhecimento do negócio, as conclusões obtidas nos procedimentos de revisão analítica e nas indagações efetuadas para suportar as avaliações do risco que efetuou. Com o conhecimento obtido pelo auditor através da utilização das técnicas de *Analytics* (e das ferramentas eletrónicas), é possível identificar as transações, processos ou contas consideráveis que acarretam um risco significativo para as demonstrações financeiras. Através de ferramentas de visualização de dados, o auditor consegue obter um melhor entendimento dos riscos significativos e riscos de fraude que, de outra forma, poderiam não ser identificados sem uma análise exaustiva. Adicionalmente, as técnicas de *Analytics* suportam o entendimento do ambiente de controlo e na identificação de riscos relacionados com controlo interno ao nível da entidade ao proporcionar um entendimento da qualidade dos registos contabilísticos dentro das demonstrações financeiras.

Os procedimentos substantivos analíticos normalmente são efetuados recorrendo às seguintes técnicas: (Marques, 2016).

1. Análise de tendências – Comparação de um conjunto de dados do período corrente com o conjunto de dados do período anterior ou com um padrão dos últimos períodos. O auditor avalia se o comportamento da rubrica está em linha com o padrão que tem seguido ao longo do exercício e se está de acordo com o comportamento esperado tendo por base os dados históricos dos períodos anteriores. As ferramentas eletrónicas que permitem a visualização das tendências ao longo do exercício são bastante úteis para este tipo de análise;

²⁴ Cfr. De acordo com o parágrafo 5 da ISA 315. A capacidade do auditor visualizar de uma forma rápida os dados da entidade através do suporte de ferramentas eletrónicas permite-lhe a utilização de critérios quantitativos (transações e saldos) e qualitativos (relações, número de processos, número de utilizadores, etc.) na identificação das contas significativas. No que respeita à tempestividade da execução dos procedimentos de auditoria, tal como descrito anteriormente, estas técnicas permitem ao auditor efetuar um processo de auditoria contínuo e assim libertar alguns procedimentos da fase final (desde que permitidos pelas Normas de Auditoria). Os procedimentos de revisão analítica podem ir desde simples comparações a análises mais complexas através da utilização, por exemplo, de técnicas estatísticas. A análise dos dados pode ser usada para apoiar a execução deste tipo de procedimentos, permitindo um nível adequado de desagregação de dados para ajudar o auditor a definir uma expectativa de que seja suficientemente rigorosa. A definição de um procedimento de revisão analítica está apenas limitado ao conjunto de dados disponíveis e à experiência e criatividade da equipa de auditoria.

2. Análise de rácios – Inclui a comparação de rácios calculados para o ano corrente com os mesmos rácios relativos ao período anterior, a uma média da indústria ou um orçamento. A análise de rácios é útil para analisar contas de ativos e passivos, bem como as contas de gastos e rendimentos. Uma conta individual balanço é difícil de prever por conta própria, mas a sua relação com outra conta é muitas vezes mais previsível (por exemplo, as contas de saldo a receber relacionados com as vendas). Os rácios podem também ser comparados ao longo do tempo ou com os rácios de outras organizações do mesmo setor de atividade;
3. Testes de razoabilidade – Consistem no cálculo de um montante previsto para o saldo da conta com base em dados não-financeiros para o período atual. Ao contrário da análise através de tendências, este procedimento analítico não se baseia em acontecimentos de períodos anteriores, mas sim em dados financeiros ou operacionais para o período de auditoria em consideração.
4. Técnicas de modelação estrutural – Utilização de ferramenta eletrónica para a construção de um modelo estatístico baseado em dados financeiros e/ou não financeiros de forma a prever o saldo de uma determinada conta. É exemplo deste tipo de análise a construção de regressões lineares. As técnicas acima identificadas diferem quer no nível de segurança obtido, quer na capacidade que apresenta em identificar distorções materialmente relevantes:

Conforme a Figura 12 o auditor deverá utilizar o seu julgamento profissional para selecionar a técnica que irá conduzir à maior probabilidade de detetar uma distorção relevante, considerando os dados disponíveis e a evidência obtida através de outros procedimentos realizados.

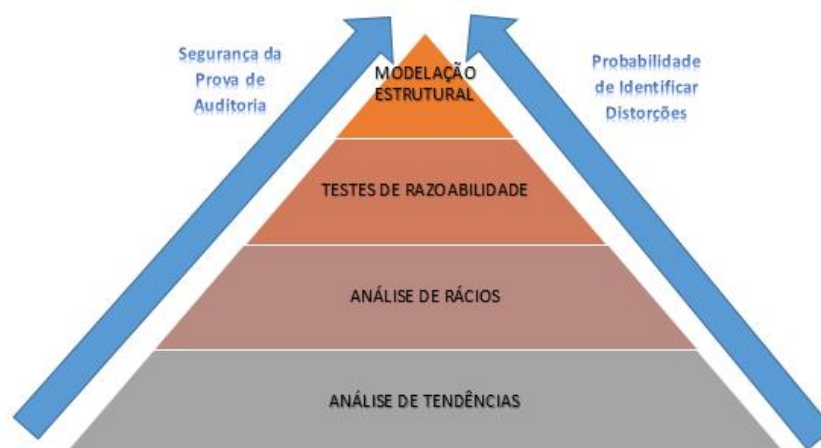


Figura 12: Hierarquia das Técnicas Utilizadas na Realização de Procedimentos Substantivos de Revisão analítica.

Fonte: Pedro Marques – OROC (2016)

2.3.2 COSO – Controlo Interno

O COSO²⁵ conclui que as organizações necessitam de orientação no momento de avaliar e melhorar o seu sistema de controlo interno. Assim, constata-se um crescimento exponencial da auditoria, pelo significativo aumento de organizações certificadas.²⁶

“O Controlo Interno é um processo, levado a cabo pelo Conselho de Administração, de Gestão, e outro pessoal, desenvolvido para fornecer uma segurança razoável no que respeita ao alcance dos objetivos relacionados com as operações, reporte, e cumprimento de leis e regulação.” (COSO, 2013)²⁷.

Com a evolução no tempo e o conseqüente crescimento das organizações²⁸ impôs-se a conceção de complexos sistemas de controlo interno, registando-se paralelamente uma evolução nos objetivos da auditoria.

²⁵ COSO – Committee of Sponsoring Organizations of the Treadway Commission - Comitê das Organizações Patrocinadoras. COSO, 2013, página 3.

²⁶ As auditorias internas podem beneficiar do método C.O.S.O. para auxiliar na identificação, controlo e análise dos riscos inerentes ao negócio, o objetivo é identificar os benefícios nos procedimentos técnicos.

²⁷ Definição de acordo com a COSO (Grupo de trabalho constituído por vários organismos internacionais, com o objetivo de definir um novo conceito de controlo interno)

²⁸ Grandes, médias e mesmo pequenas instituições costumam utilizar controlos internos para atender seus diferentes propósitos, como, por exemplo, servir de base para obter conformidade com legislações, implantar projetos de Governança Corporativa e buscar certificação por órgão acreditador, ou outros objetivos específicos.

Segundo a COSO, o controlo interno é um processo levado a cabo pelo Conselho de Administração, Direção e outros membros da organização com o objetivo de proporcionar um grau de confiança razoável²⁹ na concretização dos seguintes objetivo:

- Eficácia e eficiência dos recursos;
- Fiabilidade da informação financeira;
- Cumprimento das leis e normas estabelecidas.

O ambiente de controlo constitui o pilar de todo o sistema de controlo interno, em que fatores, tais como a integridade, ética, competência, responsabilidade, autoridade e formação do pessoal, cimentam esta componente e edificam toda a pirâmide do sistema.

O controlo interno é composto por componentes interrelacionadas, que derivam da forma como a gestão gere o negócio e de como são integradas no processo de gestão.

Para se tornar uma decisão racional, deve-se ponderar todas as alternativas existentes e identificar as suas consequências. Assim, é preciso verificar os dados obtidos e extrair as informações fundamentais. No contexto *Big Data*, os gestores deparam-se com essas informações correspondentes a um grande volume e variedade de dados desestruturados oriundos de várias fontes. (Di Martino, 2014). A grande vantagem que o *Big Data* pode trazer também para as organizações da área de auditoria é a *Analytics* para a criação de *insights* na tomada de decisão. (Gartner, 2013). Para gerir e extrair informação significativas a partir da grande quantidade de dados atual existe uma necessidade crescente de se pensar em técnicas e sistemas de *Analytics* avançados. (Esteves, Hacker & Rong, 2014).

Muitas organizações estão indecisas em adotar sistema de análise de *Big Data*. (Kwon, 2014). Logo, mesmo as organizações que não possuem um sistema de análise de grande volume de dados devem compreender o seu contexto, para que possam gerir essa variedade de dados de forma eficaz para contribuir com os processos organizacionais. Pois a informação é algo que pode ser usado para eliminar ou reduzir a incerteza. (Raskin, Taylor & Hempelmann, 2013).

²⁹ Para além de ser necessário adotar comportamentos consoante o meio envolvente, é também importante controlar continuamente o Sistema de Gestão de Controlo. Através da análise dos benefícios esperados e conseguidos, do desempenho do sistema e de possibilidades de melhoria, a organização poderá melhorar o seu funcionamento, tendo em conta todas as vantagens que daí advêm.

No caso de estudo (*Big Data & Analytics* em Auditoria) as principais competências para ir ao encontro a um controlo interno eficiente:³⁰

1. Orientação para a mudança: São competências que correspondem à monitorização da envolvente para identificar ameaças e oportunidades, assumir riscos relacionados com a mudança, obter apoio para uma grande mudança e determinar a forma de implementar uma nova iniciativa ou uma grande mudança.³¹ (Boyd, D. & Crawford, 2012).
2. Orientação para as operações: Caracterizada como os comportamentos que os líderes usam para aumentar a produtividade e reduzir os custos, tais como planeamento de atividades de curto prazo, alocação de recursos e estruturas, identificando objetivos e atribuindo prioridades;³²
3. Orientação para as pessoas: comportamentos relacionais são, de modo geral, as ações iniciadas pelos líderes para melhorar o relacionamento com os membros da sua equipa. Esses comportamentos incluem mostrar apoio e orientação, facilitar *coaching* e *mentoring*, delegar e incentivar a colaboração e trabalho em equipa.³³
4. Orientação para o contexto: caracterizada por aqueles comportamentos que os líderes apresentam quando têm de se adaptar ou lidar com a mudança no local.³⁴

De salientar que os processos de transformação organizacional são complexos e envolvem toda a estrutura organizacional. Para tal, é importante perceber qual o caminho a percorrer para chegar a bom porto. A figura 13 mostra os diversos processos para ir ao encontro de *Big Data & Analytics*.

³⁰ Para avaliar determinado objeto de auditoria, é possível analisar se a organização possui um ambiente de controlo interno eficiente (se os funcionários sabem o que deve ser feito, se sabem como fazê-lo e se querem e podem fazê-lo), se os riscos relacionados ao objeto de auditoria, relacionando-o ao objetivo estratégico da organização, está identificado, mensurado e gerenciado, se possui normas e procedimentos específicos relacionadas ao objeto de auditoria e se essas estão de acordo com os objetivos estratégicos traçados para o período.

³¹ Além dos benefícios citados, o método C.O.S.O. incrementa o universo de auditoria, exige escolhas de profissionais de auditoria com habilidade em risco, mudanças nas abordagens e na metodologia do planeamento de auditoria.

³² No decorrer de uma avaliação de riscos, cada objetivo operacional, do nível mais alto (como “dirigir uma companhia lucrativa”) ao mais baixo (como “salvaguardar caixa”), é documentado e, então, cada risco que possa prejudicar ou impedir o alcance do objetivo é identificado e priorizado (DELOITTE, 2006)

³³ Controlos internos são afetados por pessoas

³⁴ A mudança de um enfoque tradicional para o foco em riscos vem acarretando alterações significativas no escopo dos trabalhos da Auditoria Interna.

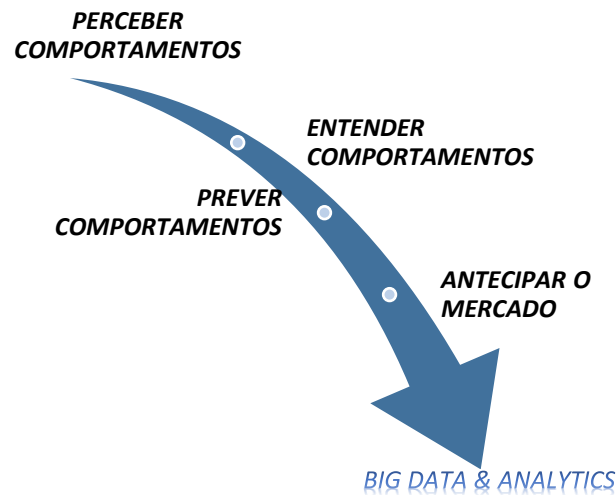


Figura 13: *Big Data & Analytics*.

Editado pela Autora

Com a utilização das ferramentas eletrónicas de auditoria, surge a necessidade de confirmar que todos os dados informáticos sobre o qual o auditor está a incidir a sua análise se encontram corretos e completos sob pena de todas as análises efetuadas não servirem de prova de auditoria. A informação recolhida para a utilização das técnicas de *Analytics* pode ser obtida por várias fontes, as quais se encontram agrupadas em 5 níveis: dados chave, balancetes auxiliares, processos, performance e dados externos, conforme se apresenta na Figura 14 que ilustra, adicionalmente, a relação entre os diversos níveis de dados, sendo o nível nuclear ocupado com os dados chave e o nível exterior com os dados externos. (Alecrim, 2013).

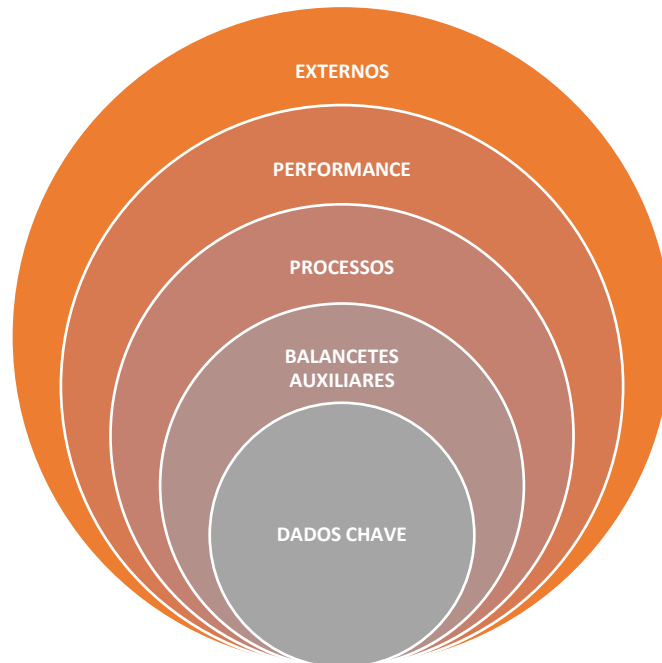


Figura 14: Fontes de dados utilizados na realização de técnicas de *Analytics*.

Fonte: Pedro Marques – OROC (2016)

Independentemente da avaliação dos sistemas de informáticos ser efetiva ou não, o auditor terá que analisar a plenitude e exatidão dos dados obtidos uma vez que os mesmos servirão de base para a prova de auditoria. Aplicação das técnicas de *Analytics* no processo de auditoria e a utilização das técnicas de *Analytics* não se limita apenas à realização de procedimentos analíticos substantivos.

Na Figura 15 mostra o controlo interno – COSO – com as suas componentes, objetivos e unidades de negócio.

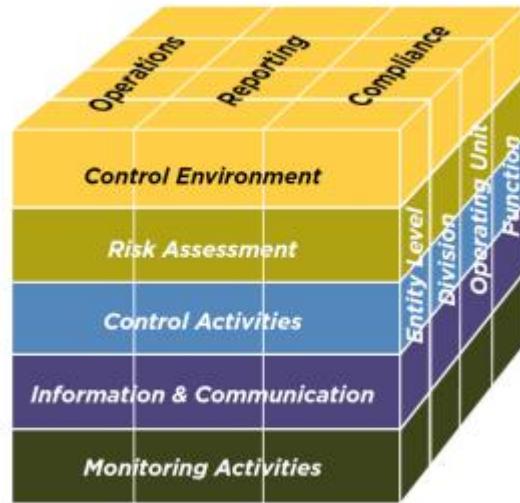


Figura 15: COSO (2013).

Fonte: <http://www.ipai.pt/index.php>

De facto, a utilização de técnicas de *Analytics* pode estar presente em praticamente todo o processo de auditoria porém, no âmbito do presente trabalho, irei apenas descrever de que forma estas técnicas podem ajudar o auditor a executar os seguintes procedimentos:

1. Conhecimento da entidade e do negócio e identificação dos riscos significativos;
2. Definição do plano de auditoria;
3. Realização de procedimentos substantivos (quer procedimentos analíticos quer testes de detalhe).

Na Figura 16, demonstra-se a forma como as técnicas de *Analytics* trazem melhorias quanto ao entendimento do negócio da entidade e no processo através do qual são identificados os riscos significativos.

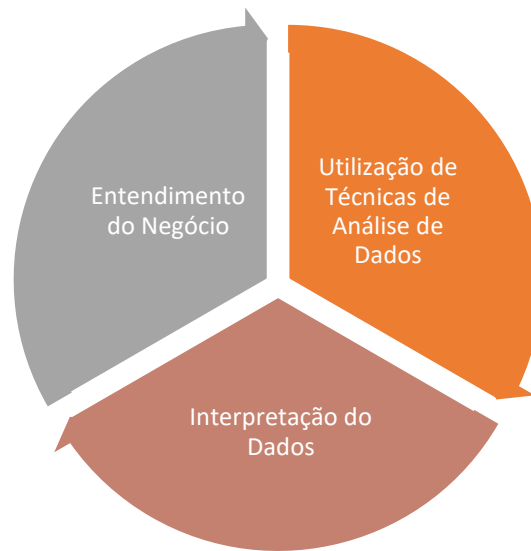


Figura 16: Ciclo da utilização da técnica de *Analytics* no conhecimento do negócio.

Fonte: Pedro Marques – OROC (2016)

A utilização da técnica de *Analytics* no conhecimento do negócio permitem sumarizar, aceder, interrogar e recalcular conjuntos completos de dados com o objetivo de um melhor entender o negócio, da estrutura e operações da entidade, assim como compreender de que forma as transações são registadas nos sistemas informáticos.³⁵ A relação entre a obtenção de conhecimento da entidade e as técnicas de *Analytics* é um ciclo, uma vez que é necessário o auditor ter um conhecimento mínimo do negócio e das operações da entidade para que possa interpretar de forma eficiente a prova de auditoria gerada.³⁶

³⁵ É necessário que os auditores mudem a sua mentalidade para abraçar um ambiente de relatórios contínuos, e, em seguida, adquiram as competências técnicas necessárias e o conhecimento da matéria para atender às exigências deste novo ambiente.

³⁶ Um dos importantes aspetos dessa nova realidade para a auditoria, consiste no facto dos procedimentos tradicionais de obtenção de provas de auditoria não poderem ser apropriados ou suficientes, já que a evidência em auditoria é gerada principalmente em formato eletrónico.

2.3.3 Como *Big Data* e *Analytics* influencia a Auditoria

A auditoria é uma das áreas que mais pode sofrer grandes transformações com as oportunidades oferecidas pelo *Big Data* e pela *Analytics*.³⁷ Os avanços das tecnologias de *Analytics*³⁸ estão a provocar alterar a forma como a auditoria é pensada e, principalmente, executada. O desafio passa pela prática da auditoria que deverá ir para além dos limites dos testes baseados em amostras, visando incluir grandes volumes de dados relevantes, nomeadamente, transações, dados de processos chave de negócio, entre outros, por meio de análises inteligentes. O objetivo é fornecer uma maior quantidade de evidências e *insights* à organização.

Existe uma série de barreiras no processo de integração de *Big Data* e *Analytics* com a Auditoria, nomeadamente, o facto de conseguir explorar os dados de uma forma eficiente, e, para tal, o uso, ainda que pouco frequente, das ferramentas de *Analytics*. Outro fator dentro deste âmbito é o investimento significativo que se faz sentir na segurança, que se pode traduzir, paradoxalmente, a uma possível maior facilidade na cedência desses dados (ou do acesso aos mesmos) a entidades desconhecidas. Para além destes fatores, acrescenta-se o facto de os auditores terem de lidar com *software* diverso, muitas vezes dentro da mesma organização, para não falar da extração dos dados exportados de *software* poderem não obedecer a um processo competente e eficaz. Outro obstáculo é o aumento na complexidade dos dados que a integração de *Big Data* com a auditoria poderá vir a trazer.

O uso de *Analytics* para produzir evidências em resposta a esses riscos³⁹ é mais difícil, no que concerne à *Analytics* com os algoritmos e regras usados para transformar dados e produzir visualizações e relatórios. De certo modo, o valor da utilização de *Analytics* em auditoria depende da importância que é dada pelos auditores para executar, ou seja, juntar “*Big Data*” com “*Analytics*” é crucial para a competitividade e rentabilidade.⁴⁰ Apesar de

³⁷ Na verdade, o pensamento por trás dessa abordagem tem suas raízes na visão estruturalista da estratégia competitiva, que trata a estrutura da indústria como dada e afetada apenas por fatores exógenos, tais como saltos tecnológicos e mudanças nas condições macroeconómicas.

³⁸ A *Analytics* tem permitido que auditores identifiquem melhores processos financeiros, fraudes e riscos operacionais, permitindo que os profissionais aprimorem sua abordagem para entregar uma auditoria mais relevante.

³⁹ Deve estar fundamentada em evidências que proporcionem aos gestores a exata mensuração das técnicas utilizadas na gestão do negócio, em que a forma de expressão do auditor, quer oral quer escrita, pode alterar a aceitação de mudanças no rumo estipulado por parte da gestão.

⁴⁰ Com os novos desafios impostos à Auditoria, a função deve estar mais atenta e diligente do que nunca. Nesse sentido, algumas ações têm que estar no radar da Auditoria no curto/médio prazo: Automatização

muitas contrariedades, aproveitando o *Big Data e Analytics* em função da auditoria para manter a gestão de risco ⁴¹, as organizações de auditoria estão a começar a sensibilizar-se para usar a tecnologia por forma a revolucionar o modo como as auditorias são realizadas.

Assim, é relevante expor as razões pelas quais se deve acreditar na integração na *Analytics* nas auditorias:

1. Acesso limitado à auditoria de dados relevantes;
2. A disponibilidade escassa de recursos qualificados e experientes para processar e analisar os dados; (Mantelero, 2014).
3. A integração atempada de análises para a auditoria é um desafio para os auditores.

Analisando os dados para produzir informações úteis é um aspeto chave do desafio e uma oportunidade para as organizações. Como pode a utilização de *Big Data & Analytics* melhorar as capacidades de auditoria de uma organização?

1. Decidir a estratégia para *Big Data & Analytics*;
2. Determinar o que é relevante;
3. Focar sobre o que vai impulsionar o valor.

De forma sumariada, neste capítulo abordou-se o facto do processo de auditoria ter sofrido alterações significativas em resultado de alterações nos sistemas de informação utilizados e pela forma como a informação está disponível, quer para o auditor quer para a gestão da organização. A evolução das tecnologias de informação⁴² abriu caminho para a introdução de novas ferramentas e metodologias de auditoria. De facto, a verificação dos dados recolhidos é fundamental nesta nova abordagem de auditoria, sob o risco de todo o trabalho ter sido suportado em dados que possam não ser os corretos, e, desta forma, colocar em causa a opinião expressa sobre as decisões.⁴³ Preferencialmente o auditor deve extrair ou acompanhar o processo de extração da informação de forma a garantir que a informação disponibilizada não foi manipulada. Outra alteração significativa na forma como o trabalho

do processo de Auditoria, através da implementação de mecanismos de Auditoria à distância e utilização de *Analytics* durante todo o ciclo de auditoria.

⁴¹ Aplicação de tecnologia para maior cobertura e visão

⁴² O Foco e mandato da Auditoria Interna tem vindo sempre a reagir perante as muitas alterações que o ambiente organizacional e regulatório sofrem. O ambiente atual é uma mistura da necessidade contínua de requisitos de conformidade com a exigência de providenciar orientações valiosas para o negócio.

⁴³ Cfr. http://recipp.ipp.pt/bitstream/10400.22/1804/1/DM_CatarinaPinheiro_2013.pdf Cada vez mais, as organizações têm necessidade de se reinventar e tomar as decisões mais assertivas contribuindo para o crescimento da organização assim como para o crescimento sustentado.

é realizado, impulsionada pela evolução das tecnologias, é a capacidade que as equipas de auditoria têm para, de uma forma rápida e sem custos significativos, reunir no mesmo projeto todas as competências necessárias para uma resposta adequada aos riscos⁴⁴ da organização.

⁴⁴ O auditor deve planear o trabalho de campo e estabelecer, a natureza, extensão, profundidade e oportunidade dos procedimentos a adotar, com vista a atingir um nível de segurança que deve proporcionar e tendo em conta a sua determinação do risco de auditoria e a sua definição dos limites de materialidade.

3 Metodologia

3.1 Método “Estudo de Caso”

Neste capítulo apresentam-se os procedimentos metodológicos utilizados para atingir o objetivo deste trabalho.

A metodologia consistiu nos seguintes passos: inicialmente foi realizada uma pesquisa bibliográfica sobre o tema (revisão de literatura), principalmente dos conteúdos disponíveis via Web. Em seguida foram analisadas algumas situações particulares em contexto de auditoria para trabalhar com *Big Data*.

O método de investigação escolhido foi o “Estudo de Caso” já que esta metodologia é ajustada quando se pretende conhecer o “como?” e o “porquê?”, com o objetivo de descrever ou analisar um fenómeno. Com este método, uma entidade é examinada, a investigação é centrada num fenómeno natural inserido num contexto real e não existem manipulações ou controlo de experiências envolvidas.

Neste “Estudo de Caso”, a entrevista é a fonte de dados utilizada, orientada por um guião de questões previamente definidas. A entrevista é utilizada para recolher dados descritivos na linguagem do entrevistado, permitindo ao investigador desenvolver intuitivamente uma ideia sobre a maneira como os sujeitos interpretam aspetos do mundo.

3.2 Estratégia da investigação

O “Estudo de Caso” é executado através de uma entrevista presencial, de onde surge matéria para a análise dos resultados decorrente da entrevista realizada. Este trabalho descreve o estudo de caso como estratégia de investigação. A compreensão deste método, enquanto estratégia de investigação emergente, pode revelar-se para investigadores que queiram estudar e inovar no campo educativo.

4 Estudo de Caso

4.1 Contexto da pesquisa

Inicialmente, para desenvolver este tema, e após a elaboração de revisão de literatura, foi feita uma análise de informação através de entrevistas. Com 5 auditores/técnicos envolvidos no processo de auditoria num contexto de *Big Data* e *Analytics*, o objetivo será elaborar entrevistas a nomes da área da auditoria de diversos sectores. A identidade dos entrevistados será estritamente confidencial. As entrevistas duraram cerca de 30 minutos cada e todos os entrevistados foram informados que a duração seria de aproximadamente esse tempo. Dada a riqueza dos dados conseguidos, considera-se este estudo de caso um dos pontos mais fortes desta pesquisa.

4.2 Desenho da entrevista

A entrevista foi elaborada com base no conhecimento adquirido na revisão de literatura, no que diz respeito ao *Big Data & Analytics*. As entrevistas foram realizadas presencialmente conforme disponibilidade do entrevistado, o que permitiu um contacto direto e informal com os entrevistados garantindo uma recolha de informação mais eficiente e ajustada ao objetivo.

A entrevista foi estruturada em 4 fases, nomeadamente:

1. Informação do respondente
 - 1.1. Função na organização
 - 1.2. Número de anos de experiência profissional
 - 1.3. Número de anos de experiência no mercado Auditoria
2. Organização onde trabalha
 - 2.1. Qual o segmento de mercado/área de atuação que atualmente trabalha/audita?

2.2.Neste momento, a organização dispõe de ferramentas para *Analytics*?

3. Contexto de utilização de *Big Data/Analytics*

3.1.A organização tem noção da dimensão dos dados gerados? Ou seja, em média quantos utilizadores têm ativos a gerarem dados diariamente?

- a. Entre 0-50 utilizadores
- b. Entre 50-100 utilizadores
- c. Entre 100-150 utilizadores
- d. Mais de 150 utilizadores

3.2.Existe controlo interno constante dessa informação?

3.3.Quem na organização faz esse trabalho de recolha, tratamento e análises de dados?

3.4.Como e com que ferramentas são extraídas os dados existentes na organização? Qual(s) o(s) *software*(s) utilizado(s)?

3.5.A organização utiliza técnicas concretas para extrair, recolher, tratar ou analisar os dados? Se sim quais?

3.6.Perante os fatores de risco a ter em conta, como é garantida e gerida a segurança e privacidade dos dados? Por exemplo, se utilizam o armazenamento de dados com proteção para identificação de pessoas, autenticação, etc. de forma a protegerem o acesso à informação

3.7.A organização redefine estratégias de negócio com base na estratégia de *Analytics*?

3.8.Dos objetivos da organização, algum está diretamente relacionado com o *Big Data* e/ou *Analytics*? Se sim quais?

3.9.Está previsto num futuro próximo a contratação de RH para este departamento? Se sim, qual a função e porquê?

4. Utilização do *Big Data* na Auditoria

4.1.Existe orçamento específico para *Analytics* em Auditoria?

4.2.Quais as razões que levam a organização a investir em *software* de *Analytics*?

4.3.Que desafios a organização encontra atualmente no tratamento e *Analytics*?
E o que se espera no futuro?

4.4. Como o *Big Data* pode ajudar a auditoria?

4.5. Considera que *Big Data* vai ser uma tendência no futuro utilizado pelos auditores?

4.6. Quais são os principais benefícios que espera do *Big Data* para a sua organização?

4.7. Considera que as tecnologias *Big Data* são eficazes para detetar e investigar riscos? De forma geral, se esta realidade permite um melhor controlo interno no futuro?

4.3 Análise de resultados

Como questão relevante do trabalho, o objetivo específico na análise dos dados que se obtiveram das entrevistas é entender os impactos do *Big Data*, o risco, o controlo e a auditoria neste contexto. Um aspeto interessante que não pode ser esquecido corresponde ao facto de as questões apresentadas terem como objetivo obter um determinado conjunto de informação respeitante a temas específicos e que estejam alinhados com base na revisão da literatura:

1. Saber como é que atualmente os auditores analisam e exportam dados;
2. Saber quais as ferramentas mais utilizadas para *Analytics*;
3. Conhecer os riscos que identificam e as melhorias a desenvolver no que diz respeito à informação;
4. Perceber como veem a tendência do *Big Data* no futuro e que benefícios antecipam.

4.3.1 Amostra do Estudo de Caso

No quadro que se segue, é detalhada a experiência profissional dos entrevistados, bem como o objetivo de ter sido selecionado este grupo de entrevistados.

Tabela 2: Amostra do Estudo de Caso

ID	Segmento de Mercado	Experiência Profissional (anos)	Experiência Auditoria (anos)	N.º de Utilizadores ativos	Objetivo
1	Formação Profissional	14	14	> 150	Perceber a opinião dos especialistas de Auditoria
2	Serviços de Saúde	13	6	> 150	
3	Indústria da construção, moldes e serviços	17	10	Entre 0-50	
4	Serviços de concessões rodoviárias	26	15	> 150	
5	<i>Software de Business Intelligence</i>	16	16	> 150	
6	Retalho Têxtil	13	10	> 150	
7	Banca e Seguros, Telecomunicações e Retalho	17	15	> 150	

Fonte: Elaboração Própria

4.3.2 Controlo dos dados gerados e responsáveis

Na pergunta para apurar se existe o controlo interno constante da informação a que estes entrevistados estão diariamente sujeitos, a maioria relata que ainda não existe muito controlo interno. No entanto, para outro entrevistado ele próprio é o profissional que quando solicita dados eles já vêm tratados, analisa e procura evidências de controlos e monitorização dos mesmos. Interessante é também haver organizações em que esta responsabilidade está sobre o alcance de equipamentos, que gerem as ações dos utilizadores. Grosso de modo, ainda não existe uma cultura organizacional de controlo de informação. É um percurso que está a ser percorrido. Já alguns que fazem e fazem muito bem, mas a maior parte ainda não tem o verdadeiro controlo da informação que gera e que utiliza. Muitas vezes não têm noção da informação que está a utilizar.

4.3.3 Responsáveis pela *Analytics*

À questão de quem na organização é que faz o trabalho de recolha, tratamento e análises de dados foi respondido da seguinte forma: a maior parte das entrevistas demonstram que este género de trabalho é dividido entre os diversos departamentos existentes da organização, nomeadamente, o departamento de qualidade, departamento financeiro, departamento produção, entre outros. Contudo, em organizações de menor dimensão ainda é a gerência que faz diretamente a análise. O grande desafio passa pelo conhecimento dos modelos de dados, pelo que têm de ser assertivos e internamente fazem o apoio operacional.

Importa ter noção de que existem dois tipos de organizações: a) as pequenas e médias organizações e b) as grandes organizações. As pequenas e médias organizações ainda não têm um grupo definido para esse tipo de trabalho, sendo um trabalho que, normalmente, é dividido entre os diversos departamentos, financeiro, planeamento, entre outros. Nas grandes organizações já se vê, de facto, uma equipa formada para este género de trabalho. Na maioria dos casos apresentados, numa pré-análise existe uma equipa de serviços informáticos que garante que os dados estão a ser integrados com segurança, nomeadamente os dados automáticos que resultam de vendas, contagem, fluxos de stocks, entre outros, e, para além disso, existem os dados que são inseridos manualmente, mais associados aos dados qualitativos e relativamente à formação interna. Nestes casos são os colaboradores de cada área que inserem as informações na ferramenta, para depois poderem ser tratados também por eles, de acordo com os pedidos de análise específicos feitos pela gestão.

4.3.4 Ferramentas de Extração de Dados

Sobre quais as ferramentas que são utilizadas para extrair os dados existentes na organização comprova-se, da amostra, que o Excel ainda é a ferramenta mais utilizada pelas organizações independentemente da dimensão das mesmas. Esta é a principal ferramenta para extrair os dados e para os tratar. Para além disso o Access é também utilizado para esse efeito.

Grande parte das vezes a informação já vem tratada e depois será analisada através do Auditor, confrontando dados de várias fontes, e procedendo à verificação sobre a

existência de discrepâncias, se há possibilidade de aumentar a produção, se está a ser utilizada a capacidade instalada, entre outros. Curiosamente, também foi registado que ainda se utiliza o formato de desenvolvimento do próprio *software* com base nas necessidades existentes, ferramentas próprias e customizadas, muito embora se utilizem também ferramentas convencionais de análise de indicadores de desempenho de processos, indicadores ambientais, indicadores financeiros, marketing de entre outros. Existe *software* tradicional que origina informação estruturada, como por exemplo o SAP e o SAGE, e *software* que permite uma maior autonomia para os utilizadores desenvolverem o seu trabalho. Assim, compete aos sistemas, através da utilização de rotinas informáticas automatizadas, a realização do processamento automático de extração de dados.

Foi possível averiguar que os serviços de suporte informático utilizam macros para recolha e tratamento dos dados os quais estão inseridos no sistema interno. Os colaboradores dispõem ainda de diferentes permissões, consoante as suas funções, que permitem de forma automática extrair informação específica, existente no *software*.

4.3.5 Técnicas concretas para extrair, recolher, tratar ou analisar os dados

À pergunta se a organização utiliza técnicas concretas para extrair, recolher, tratar ou analisar os dados, a totalidade das respostas demonstra que não existem técnicas concretas. Isto significa que o registo é individual, feito por projeto, caso a caso, com base na informação relevante para a organização.

Um outro motivo de reflexão, também aqui a técnica ainda passa muito pelas macros pré-definidas. Posteriormente, os dados são analisados individualmente e transformados em relatórios tipificados.

4.3.6 Gestão de risco e privacidade dos dados

A análise e discussão deste assunto aponta para que ainda não há gestão da segurança e privacidade dos dados na maior parte das situações: embora existam algumas formas de proteção de dados, estas ainda não constituem uma prioridade no meio organizacional.

Perante os fatores de risco a ter em conta, nomeadamente, como é garantida e gerida a segurança e a privacidade dos dados, podendo discriminar a evidência dos seguintes controlos:

1. Login com acessos limitados, através de autenticação;
2. Autenticação através de *password*;
3. Utilizador com segregação de acessos em função das funções desenvolvidas na instituição;
4. Confidencialidade de clientes;
5. Dados agregados sobre os clientes para tratamento de informação.

Nos casos presentes e de forma geral, a segurança é uma área estritamente garantida pelos serviços de informática, quando nos referimos aos dados informáticos, cabendo-lhes gerir cada perfil de colaborador, por forma a atribuir ou retirar permissões de acesso aos dados conforme seja necessário. Quanto aos dados e informações físicas que são usadas muitas vezes nas auditorias como evidências, estas encontram-se guardadas em arquivo com acesso restrito e são da responsabilidade de cada departamento.

4.3.7 Estratégia de Negócio Versus Estratégia de *Analytics*

Em decorrência deste enfoque, existe consenso entre os entrevistados pelo lado positivo. Já é possível perceber a existência de estratégias com base na informação. Verifica-se na amostra que já começam a existir estratégias com base na informação, já se tomam decisões com base em análises (e não com o tradicional *feeling*): embora ainda não haja 100% de confiança, a *Analytics* começa a ter a sua real importância.

Nesta matéria é preciso ter em conta que os dados recolhidos em cada auditoria, são sempre reportados e apresentados equipa a equipa, com informação de áreas de intervenção necessárias, sendo que cada Diretor tem a responsabilidade de intervir e implementar na sua área, todas as ações necessárias com vista à melhoria contínua. De referir que este facto vai pressionar a estrutura no sentido de adaptar formas de fazer e estratégias, sendo estas questões, então, decididas em Conselho de Gestão quando estão em causa alterações significativas na estrutura.

4.3.8 Objetivos relacionados com o *Big Data* e/ou *Analytics*

Nesta questão, sobre se nos objetivos da organização, algum está diretamente relacionado com o *Big Data* e/ou *Analytics*, as respostas dividiram-se: se, por um lado, a evolução das diversas plataformas existentes para uma solução única integrada é o objetivo primordial, outras preocupam-se apenas em cumprir com a legislação.

É de notar que o *Big Data* começa a aparecer nos meios de comunicação e que as organizações já começam a ter interesse e perspetivas nesse sentido. É um conceito já muito falado na teoria e que aponta para a exploração para ir ao encontro dos objetivos da organização no futuro.

4.3.9 Recursos Humanos versus *Analytics*

Na pergunta em análise, se está previsto num futuro próximo a contratação de RH para este departamento, os resultados apontam para a não existência de uma perspetiva de contratação de profissionais nesta área. Numa outra entrevista foi mencionado que já existe um departamento que assegura a evolução das necessidades quanto a esta área. Otimizar rotinas de dados é otimização recursos, uma questão que está muito dependente do tamanho da organização.

4.3.10 Orçamento versus *Analytics*

Sobre se existe orçamento específico para *Analytics* em Auditoria, as respostas dividem-se: há sectores onde esta rubrica já se encontra presente nos orçamentos, mas ainda há quem não a considere uma prioridade. Pode-se considerar que realmente que a *Analytics* poderá vir a reconhecer importância nos orçamentos futuros.

4.3.11 Investimento em *Software* de *Analytics*

No que diz respeito ao investimento e sobre quais as razões que levam a organização a investir em *software* de *Analytics*, as respostas também em muito se assemelham, nomeadamente:

1. Maior rapidez;
2. Qualidade de informação;
3. Identificar de erros, não conformidades;
4. Necessidade de acompanhar clientes;
5. Identificar formas de evolução/crescimento.

É primordial perceber que investimento se traduz no retorno do investimento, para explorar mais dados, e ir ao mais ínfimo pormenor, e que este é o segredo do sucesso por isso esta é a mais-valia das ferramentas para o *reporting*.

4.3.12 Desafios no tratamento de *Analytics*

À questão sobre quais os desafios que a organização encontra atualmente no tratamento e *Analytics* as respostas incidem sobre:

1. A falta de homogeneidade dos dados obtidos;
2. Informação relevante para a sua atuação no mercado;
3. Qualidade dos dados (nem sempre são estruturados);
4. Necessidades de estarem alinhados consoante as expectativas;
5. Capacidade do equipamento;
6. Privacidade dos dados.

A imensidão de dados produzidos todos os dias e a necessidade constante de mudança, por vezes trazem desafios de tempo, pois a análise detalhada, não sendo automatizada e requerendo reuniões entre envolvidos, obrigam a utilização de tempo, e, assim, a resposta e a mudança desejadas ficam dependentes desta disponibilidade, e não são imediatas.

4.3.13 *Big Data* Versus Auditoria

Nas questões sobre a forma como o *Big Data* pode ajudar a auditoria, a resposta foi, maioritariamente, que é considerada uma mais-valia, nomeadamente para agilizar os tempos de auditoria, otimizar os resultados esperados e garantir acesso à informação de forma

expedita, endereçando assim as questões sobre: multiplicidade, diversidade, integridade, volume e transferência de dados.

Quanto mais informação se puder recolher melhor. O problema está nos detalhes e com *Big Data* consegue-se mais capacidade de explorar esses detalhes de elevada importância.

Através da geração de padrões de utilização permitirá construir cenários desviantes que, ao concretizarem-se, ativarão sistemas de alerta precoce permitindo a intervenção da instituição antes do cenário de alerta se concretizar totalmente.

4.3.14 *Big Data* a tendência do futuro

À questão sobre se o *Big Data* será uma tendência no futuro e se será utilizado pelos auditores, ainda subsistem algumas dúvidas relativamente a este assunto e as opiniões dividem-se. Será, de facto, a forma mais eficiente de otimizar a ferramenta e monitorização continua para além de que permite ter resultados de exceção (comparabilidade de dados).

4.3.15 Benefícios do *Big Data*

À pergunta se considera que as tecnologias *Big Data* são eficazes para detetar e investigar riscos, posso destacar as seguintes:

1. Tratamento sistémico de grande volume de dados permite obter resultados mais rápidos e identificar atempadamente erros/anomalias/desvios;
2. Redução de custos;
3. Rapidez na tomada de decisão;
4. Identificação de novos serviços/soluções/produtos;
5. Sistematizado – facilita a “ponte” e é mais eficiente e mais rápido;
6. Repositório único dentro da organização;
7. Integrar dados estruturados e não estruturados;
8. Possibilidade de melhorar as capacidades de análise existentes atualmente.

4.3.16 *Big Data* versus Riscos

Se é considerado que as tecnologias *Big Data* são eficazes para detetar e investigar riscos, a grande maioria respondeu sim a esta questão positivamente. A possibilidade de identificar novas soluções cria proteção (do lado económico) e minimiza riscos, ou seja, otimiza rotinas para agir em tempo útil.

Estamos a meio caminho de um percurso e ainda há um longo caminho a percorrer no que toca à gestão e ao *Governance* dos dados, podendo ser melhorados e explorados no que respeita às análises e previsões, bem como na componente estratégica. Ainda assim, é, claramente, um vetor a explorar num futuro próximo.

4.4 Discussão dos Resultados

As entrevistas realizadas permitiram recolher um conjunto de considerações substancialmente importantes para o objetivo final deste trabalho. De acordo com o acima exposto, é possível fazer umas breves conclusões.

O estudo de caso vem confirmar a utilização de diversas técnicas e *software* com o objetivo de obter conhecimento, tendências e comportamentos do utilizador. Em todos os casos, é possível constatar que existe uma equipa dedicada na área da gestão e *Analytics*, sendo expectável que venha a crescer nos próximos meses, porque há realmente a necessidade que isso aconteça. No entanto, torna-se importante refletir que ainda se considera que ainda não existe nas organizações incutido o conceito de cultura orientada para os dados.

O investimento nesta área ainda não é o que se espera para o futuro. Contudo, os tempos modernos requerem que se preveja um investimento francamente significativo a longo prazo uma vez que o orçamento global da organização ainda não está a direcionar para a temática dos dados, recursos humanos especialistas e tecnologia específica. Apesar deste facto nota-se uma crescente sensibilização para que tal venha a suceder numa realidade muito próxima. No seguimento do que foi escrito anteriormente no ponto 2.2.2. “A transformação de Dados” é fulcral perceber que é um investimento no *Big Data* e *Analytics* que se destina à melhoria em termos de eficácia ou eficiência de negócio.

Na minha perspetiva, e da forma como ainda se aborda o tema, que ainda é novidade, o acompanhamento ou o controlo interno nesta vertente fará mais sentido que, no futuro, venha a decorrer em tempo real, para que se possa acompanhar as principais métricas. No que à responsabilidade de *Analytics* diz respeito, espera-se que seja uma equipa de elementos a analisar com mais detalhe essa informação: muito embora tal já aconteça nas grandes organizações, esta função vai criar postos de trabalho para além de pessoas especializadas no assunto que irão ter um certo tipo de sensibilidade para tratar dos dados com muito mais cuidado do que já se faz atualmente, e prevê-se a existência de técnicos com mais competências de *Analytics* e focados em análises mais aprofundadas, tal como vimos no ponto 2.1.2. “O que é o *Big Data*?” 2. É necessário a Gestão de Talentos: devem atrair-se profissionais qualificados para trabalhar com grandes quantidades de informação.

Tive oportunidade de concluir, e respondendo também a uma das questões a que propus ser o objetivo final do estudo de caso, saber como é que atualmente os auditores analisam e exportam dados, e tendo em conta o mercado de ferramentas (*softwares*) que existem hoje em dia, são as ferramentas tradicionais que ganham destaque no nosso mercado nacional pelo que se destaque o ainda tradicional Excel. O funil de dados é muito semelhante, isto é, extração, segmentação, tratamento e *Analytics*. Tal como vimos no ponto 2.2.3. “Vantagem Competitiva” existe uma variedade de ferramentas (*software*) que são usadas para executar *Analytics* e ajudam a identificar áreas chave de risco.

No que respeita às técnicas utilizadas, espera-se que as organizações acompanhem a evolução da tecnologia, o que não se conclui que seja uma tendência de hoje, talvez por definição de prioridades, mas o que se espera é que as técnicas analíticas façam parte do trabalho do auditor para prever e retirar conclusões acerca dos dados. Também responde à segunda questão do objetivo final sobre quais as ferramentas mais utilizadas para *Analytics*, tendo em linha de conta o ponto 2.3. “Auditoria – Requisitos Mínimos de Controlo”, é necessário o uso de ferramentas e técnicas de *Analytics* especializadas, pois as atuais serão ineficientes.

Outro aspeto fulcral nesta análise foi a gestão da segurança e privacidade de dados, pois este é dos pontos mais desafiantes deste mercado. É notável a importância e preocupação dada à segurança dos dados nos tempos atuais, que garantem a integridade, privacidade e a autenticidade dos dados do utilizador. E aqui sim já faz parte das prioridades

de segurança de informação. O caminho é ter um sistema mais eficiente possível para salvaguardar a segurança de toda a estrutura organizacional. O armazenamento de dados, a proteção para identificação de pessoas, autenticação, comunicação encriptada e *firewall*. No seguimento do ponto 2.3.1. “Riscos do *Big Data*” a preocupação com a segurança da informação é crescente, compatível com a necessidade de proteger dados confidenciais de organizações e pessoas.

Outra questão consiste no reconhecimento de que, pelo menos dos dados que obtive, a estratégia de negócio não passa pelo *feeling* que tradicionalmente era usado como o impulso para qualquer decisão. Hoje em dia, e tal parece ser um bom indicador toma-se em conta os dados exportados, para além de que as estratégias são redimensionadas tantas vezes quantas as necessárias para irem ao encontro dos resultados esperados, sendo que o modelo de negócio é focado nos dados. Tal como mencionado no ponto 2.3.3. “Como *Big Data* e *Analytics* influencia a Auditoria” o pensamento por trás desta abordagem tem as suas raízes na visão da estratégia competitiva.

Há preocupação de investimento em matéria de dados, mas ainda não tem a importância esperada. Na minha perspetiva, este futuro investimento será dividido em recursos humanos e *software* pois aproveitar estes recursos será uma mais-valia para melhor conhecimento e *insights* que correspondem aos interesses da grande parte das organizações.

Em termos de desafios, muitos são eles que estão “à porta” de quem quer evoluir neste âmbito, saber escalar o negocio da melhor forma, e com isso pode-se contar com o continuo aumento de dados. No que se refere ao ponto 2.3.2. “COSO – Controlo Interno” o desafio das organizações é extrair informações estruturadas, logo, as organizações que aprenderem a tirar proveito do *Big Data* poderão utilizar informação em tempo real.

Em relação aos desafios que se avizinham neste tema há destaque para a falta de homogeneidade dos dados obtidos, informação relevante para a sua atuação no mercado, qualidade dos dados (isso confirma-se quando é constatado que nem sempre são estruturados), a falta de alinhamento dos mesmos consoante as expectativas, capacidade do equipamento e privacidade dos dados. Por outro lado, consideram-se como benefícios o tratamento sistémico de grandes volumes de dados, o que permite obter resultados mais rápidos e identificar atempadamente erros/anomalias/desvios, redução de custos, rapidez na tomada de decisão, identificação de novos serviços/soluções/produtos, sistematizado –

facilita na ponte entre a informação e o trabalho do auditor, é mais eficiente e mais rápido, havendo um repositório único dentro da organização e integrando dados estruturados e não estruturados. Este ponto também vem responder à terceira questão do objetivo final sobre quais os riscos e as melhorias a desenvolver no que diz respeito à informação. Tal como já foi referido no ponto 2.3.1 “Riscos do *Big Data*” entre os principais benefícios que as organizações nacionais encontram na transformação digital do negócio destaca-se a eficiência operacional, um melhor relacionamento com os clientes e a criação de novas fontes de receita.

Por último e não menos importante, realçar que maioritariamente nas organizações dos entrevistados existem mais de 150 utilizadores ativos o que se conclui que os dados transmitidos pela organização transmitem a relevância que a aplicação assume, com um número de utilizadores bastante considerável, sendo que aqui já se trata efetivamente de *Big Data*. Há efetivamente consciência neste tema, e nota-se o empenho em retirar o melhor conhecimento possível dos dados que são rececionados conforme as disponibilidades/recursos que cada dimensão Empresarial pode assumir. Em função da auditoria ainda não se tem a perceção se será uma tendência do futuro neste ramo, mas será certamente a forma mais eficiente de otimizar as ferramentas e na monitorização continua para além de que permite ter resultados de exceção (comparabilidade de dados). E aqui responde à quarta e última questão que propus como objetivos como é que veem a tendência do *Big Data* no futuro e que benefícios podem colher.

Após o trabalho das entrevistas, e ao confrontar os referenciais teóricos com a prática, constata-se que os dados são um pilar fundamental da organização. No entanto, posso constatar que da revisão de literatura que fiz, este é um tema atual ao nível mundial, mas que ainda está longe de ser o dia-a-dia da maior parte das organizações a nível nacional. Estamos a caminhar nesse sentido mas as mudanças ainda não se fazem notar tanto quanto se esperava.

5 Conclusões e Trabalho Futuro

O conceito de *Big Data* não sobrevive sem tecnologia. Mas a afirmação crescente do conceito junto das organizações acontece porque o *Big Data* só existe e faz sentido se for usado para tomar melhores decisões, com maior probabilidade de êxito e minimizando o risco. Este estudo permitiu-me tirar uma breve conclusão sobre o assunto, que concerne no seguinte: o *Big Data* é um problema da informática pelos riscos do poder da informação, mas é, certamente, uma solução para as áreas de negócio que pode vir a fazer toda a diferença. É indiscutível que o volume de dados está em constante crescimento no seio das organizações em todas as suas direções, divisões e departamentos. Saber gerir é, cada vez mais, um marco a cumprir das organizações. Deve-se fazer um balanço dos dados que se possui, deixar que prevaleçam as necessidades dos negócios, reavaliar a infraestrutura, estruturar os dados e preparar a equipa.

O *Big Data* é a solução que contribui de forma relevante para o futuro das organizações, garantindo vantagem competitiva às organizações. Com base na investigação realizada neste trabalho, verifica-se que, os avanços tecnológicos, os quais envolvem direta ou indiretamente a sociedade, encontram-se de tal forma desenvolvidos que não há como fugir do fenómeno *Big Data*. Torna-se assim um dos principais conceitos nas tecnologias de informação para auditoria embora ainda não sendo uma realidade que se avizinha de implementação imediata. A auditoria, os profissionais que controlam o risco de informação e dados, têm um papel fundamental para apoiarem nesta fase em consonância com a utilização apropriada dos dados organizacionais. De forma geral, a tecnologia permite analisar qualquer tipo de informação digital em tempo real quer para ajudar nas decisões das organizações, quer para melhorar a competitividade, produtividade e inovação.

Para a realização deste trabalho foram consultados livros que abordam o assunto *Big Data* e artigos de investigação. Apresentou uma síntese da questão que está a vincar de forma relevante a realidade de todo o ambiente organizacional e social, aspeto que nos leva às dimensões do risco inerente à questão do impacto que causaria a utilização da análise de *Big Data*. O objetivo desse trabalho é o de identificar o ambiente *Big Data* com requisitos de controlo que se pode vislumbrar num futuro próximo, tornando-se relevante saber como

colher frutos do “Fenómeno *Big Data*” e beneficiar do impacto positivo que a *Analytics* digitais pode oferecer.

Os resultados obtidos através das análises realizadas permitem uma conjugação de informação com origem em dados estruturados e dados não estruturados. O Estudo de Caso apresentado serve, também para mostrar que ainda existe muito trabalho a realizar nesta área, pois, para uma correta utilização de dados não estruturados, é muito importante perceber, em primeiro lugar, o que esperamos obter com a análise dos dados e, em segundo, perceber o valor que esses dados não estruturados podem ter e como podem ajudar os processos da organização. As aplicações que fizeram uso do *Big Data* em conjunto com modelos de realidade aumentada poderão oferecer as melhores condições de efetividade na análise de cenários de elevada complexidade, e portanto, de competitividade e, quando envolverem o ser humano num processo colaborativo. Estes indícios nos levam a refletir no sentido de para aproveitar a janela de oportunidades do *Big Data* e se beneficiar do seu potencial competitivo, as organizações devem ser proactivas sendo necessário reconhecer os fenómenos por detrás do *Big Data*, estudá-los em profundidade e avaliar seriamente os seus impactos.

“O Big Data representa a possibilidade de re-imaginar o nosso mundo, percecionarmos novos sinais, e mudar a forma como experienciamos as nossas comunidades, os nossos locais de trabalho e as nossas vidas pessoais.”, (The Human Face of *Big Data*, 2016). Seguindo uma fundamentação teórica, acredito, que para que *Big Data* realmente se torne uma força de mudança no mundo dos negócios, tal como aconteceu há décadas com os computadores pessoais, com a Internet e com a comunicação móvel, terão de surgir ferramentas poderosas, acessíveis e gratuitas, sem necessidade de aprender outras tantas ferramentas e linguagens do *Big Data*. Os projetos de *Analytics* tendem a ser dispendiosos, quer em tempo, quer em custos, pois não é fácil o cliente perceber o que realmente pretende sem antes ver alguns dos dados. Uma das premissas básicas deste paradigma, diz que a consciência sobre a evolução das soluções tecnológicas é o ponto de partida para que esse futuro esteja cada vez mais próximo. Usufruir dessa inovação significa melhorar a gestão de recursos, por meio da deteção e correção tempestivas de deficiências de controlos internos, erros e fraudes, bem como do aperfeiçoamento da gestão de riscos.

Como resultado, é evidenciado, então, o importante papel dos órgãos de controlo que passem a dispor de ferramentas cada vez mais eficientes e eficazes para assegurar a boa gestão dos recursos.

Importa ainda destacar, que, do trabalho realizado e como resultado da revisão da literatura, foi publicado um artigo científico na 11.^a CISTI, Conferência Ibérica em Sistemas e Tecnologias de Informação, que decorreu na Gran Canária, em Junho de 2016. O artigo encontra-se indexado nos motores de busca: EI, IEEE Xplore, INSPEC, ISI, SCOPUS e Google Scholar.

5.1 Trabalho futuro

Os investigadores têm alertado para a escassez de estudos refletindo as implicações da utilização das ferramentas de *Big Data & Analytics* em questões organizacionais, culturais e sociais mais abrangentes. (Boyd, & Crawford, 2012). Especialmente em Portugal, o volume de pesquisas académicas em *Big Data* ainda é pequeno.

Como oportunidade para futuros estudos deixo alguns pontos relevantes ter em conta:

1. Estudos Multidisciplinares, sobre *Big Data & Analytics*, discutindo contextos, aplicações e questões específicas das áreas, sob óticas mais abrangentes.
2. Estudo de padrões metodológicos: diversos autores alertam sobre a importância da participação académica nas discussões mais amplas sobre a aplicação de *Big Data* para definir ou, pelo menos, influenciar os padrões metodológicos a serem adotados no desenvolvimento das ferramentas de *Big Data*.
3. O estudo sobre as questões éticas, de segurança e de acesso aos dados é também uma questão para o estudo futuro, já que a evolução das possibilidades tecnológicas viabiliza uma série de transações de captura e acesso a dados, algo inimaginável há alguns anos. No entanto, o facto de ser possível não significa que seja ético e é necessário discutir os sistemas que estão direccionados e regulando estas práticas.
4. Diversos autores alertam para a atual falta de profissionais especialistas em *Big Data & Analytics* no mercado e seu iminente agravamento devido ao aumento da

adoção pelas organizações sem correspondência equivalente na formação de profissionais. (Letouzé, 2012).

5. Avaliar os impactos políticos e jurídicos do *Big Data*.

6. Avaliar fragilidades tecnológicas e o seu impacto numa sociedade cada vez mais tecnologicamente dependente.

7. Avaliar as técnicas e ferramentas disponíveis para a criação e gestão de dados, na perspetiva de segurança (backup).

8. Avaliar a vida útil e analisar os aspetos tecnológicos e económicos da obsolescência acelerada dos dispositivos de armazenamento de dados, na perspetiva do *Big Data*;

9. Analisar as técnicas e mecanismos de armazenamento e recuperação de dados, na perspetiva do *Big Data*.

Estes são apenas alguns exemplos na partilha de ideias de investigação que possam inspirar outros autores da área capazes de realizar o potencial de valor deste fenómeno para a ciência e para a prática. Propõe-se pesquisas em organizações que adotem diferentes tecnologias de *Analytics* para o contexto de *Big Data*, de todas as áreas e não somente da área de TI, buscando melhorias nos processos.

Estas são algumas das numerosas razões pelas quais se deve explorar o tema e se mostra necessário refletir no sentido de que deve haver mais conhecimento e partilha do mesmo.

5.2 Limitações do estudo

Pelo exposto, desde já se pode adiantar que a tarefa que se pretendeu desenvolver não foi fácil. Muito menos capaz de estabelecer conceitos definitivos e consensuais. Primeiro porque é um tema que está na ordem do dia, mas do qual ainda muito pouco se fala em

Portugal. Existem poucos estudos neste âmbito o que dificultou o estudo a nível nacional e a comparação a nível mundial.

Estabeleceu-se, então, um conjunto de questões às quais se procurou responder no decurso deste trabalho, para que, de alguma forma, fosse facilitado o trabalho de discussão ao nível do estudo de caso. Outro fator importante, no que ao estudo de caso diz respeito: a grande maioria dos entrevistados tem muito pouco contacto com este tema, o que, de alguma forma, permite concluir que ainda há muito a fazer e a divulgar no seio das organizações.

Referências bibliográficas

Agrawal, Divyakant, Sudipto Das, and Amr El Abbadi. 2011. “Big Data and Cloud Computing: Current State and Future Opportunities.” *4th International Conference on Extending Database Technology: Advances in Database Technology, EDBT 2011*, 530–533.

Alecrim, Emerson. 2013. “O Que É Big Data?” *Infowester*. <https://www.infowester.com/big-data.php>. Acesso em 7 de Julho de 2016.

Amaral, Fernando. 2014. “Big Data e Seu Impacto Nas Auditorias Internas.” *Datascientist*. <http://www.datascientist.com.br/2014/08/big-data-e-seu-impacto-nas-auditorias.html>. Acesso em 8 de Junho 2016.

Barlow, Mike. 2013. *Real-Time Big Data Analytics: Emerging Architecture*. O’Reilly Media, Inc

Beulke, Dave. 2011. “Big Data Impacts Data Management: The 5 Vs of Big Data.” *DaveBeulke.com*. <http://davebeulke.com/big-data-impacts-data-management-the-five-vs-of-big-data/>.

Bigdata.black. 2012. “What Is Big Data?” *Bigdata.black*. <http://bigdata.black/featured/what-is-big-data/>. Acesso em 13 de Julho de 2016.

Birchfield, Reg. 2013. “Big Data: Management’s New Big gun.New Zealand Management.” *Questia*. <https://www.questia.com/magazine/1G1-346337032/nzim-big-data-management-s-new-big-gun-big-data>.

Boyd, Dannah, and Kate Crawford. 2012. “Critical Questions for Big Data.” *Critical Questions for Big Data: Provocations for a Cultural, Techonological, and Scholarly Phenomenon. Information* 15 (5): 662–679

Bughin, J, Michael Chui, and J Manyika. 2010. “Clouds, Big Data, and Smart Assets: Ten Tech-Enabled Business Trends to Watch.” *McKinsey Quarterly*, 75–86.

Candido, Petronio, 2015, O Processo da Informação. Disponível em <http://pt.slideshare.net/petroniocandido/big-data-e-data-science-admirvel-mundo-novo-iv-sic-ifnmg> Acesso em 5 de Agosto de 2016.

Chen, Hsinchun, Roger H L Chiang, and Veda C. Storey. 2012. *Business Intelligence and Analytics: From Big Data to Big Impact. Management Information Systems Quarterly*. Vol. 36, 1165 – 1188.

Claro, Daniela Barreiro, 2012. Tópicos Avançados em Banco de Dados (Conteúdo Programático: Dados Estruturados, Dados Semi-Estruturados e Dados não-estruturados. Ciência da Computação – Ufba – Unifacs. Comstor, 2013 - Riscos *Big Data*. Disponível em: <http://blogbrasil.comstor.com/bid/305729/os-riscos-do-big-data-e-a-invas-o-de-privacidade-nos-eua> Acesso em 14 de Junho de 2016.

Davenport, Thomas H. 2012. *Enterprise Analytics Optimize Performance, Process, and Decisions Through Big Data*.

Deloitte. 2016. “Oportunidades Emergentes O Exercício de Navegar Em Busca de Oceanos Azuis.” <https://www2.deloitte.com/content/dam/Deloitte/br/Documents/about-deloitte/MundoCorporativo51.pdf> Acesso em 30 de Agosto de 2016.

Dumbill, Edd. 2012. *Planning for Big Data*. O'Reilly Media, Inc. em <https://www.oreilly.com/ideas/what-is-big-data> Acesso em 27 de Julho de 2016.

Eaton, Chris, Tom Deutsch, Dirk Deroos, and George Lapis. 2012. *Understanding Big Data Analytics for Enterprise Class Hadoop and Streaming Data*. *Journal of Chemical Information and Modeling*. Vol. 53

Esteves, Rui Máximo, Hacker, Thomas e Rong, Chunming, 2014. A new approach for accurate distributed cluster analysis for *Big Data*: competitive K-Means, *Int. J. Big Data Intelligence*, Vol. 1, pp. 50-64.

Franks, Bill. 2012. “Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics - Finding Opportunities in Huge Data Streams with Advanced Analytics,” Willey, ISBN: 978-1-118-20878-6

Gantz, John, and Reinsel, David. 2011, “Extracting value from caos,” IDC Go-to Market Services, June, pp. 1–12.

Gantz, John, and David Reinsel. 2012. “The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East.” *Idc* 2007 1–16.

Gartner. Big Data. www.gartner.com, IT Glossary. Disponível em: <http://www.gartner.com/it-glossary/big-data/>. Acesso em: 21 de Julho 2016.

Gonçalves, Rui. 2015 “A importância da Gestão de Risco e o ISO31000” *Business Analytics*. <http://businessanalytics.pt/iso31000/>

Gouveia, Luís Borges, and Ranito, João. 2004. “Sistemas de Informação de Apoio À Gestão.”, Livro VII - Colecção Inovação e Governância nas autarquias, SPI – Principia, ISBN: 972 8589 43 3

IFAC. “ISA - Parágrafo 5 da ISA 315.”

ISACA, 2011. “Data Analytics — A Practical Approach”, White Paper, ISACA.Itinsight.pt. 2016. “Empresas Portuguesas Têm Pouca Maturidade No Que Respeita À Segurança.” *Itinsight.pt*. <http://www.itinsight.pt/news/seguranca/empresas-portuguesas-tem-pouca-maturidade-no-que-respeita-a-seguranca>. Acesso em 15 de Agosto de 2016.

Koops, Bert-Jaap. 2014. “The Trouble with European Data Protection Law.” *International Data Privacy Law* 4 (4): 250–261.

Kuenkaikaew, Siripan. 2013. “Preditive Audit Analytics Evolving To A New Era.”

Kwon, Ohbyung, Namyoon Lee, and Bongsik Shin. 2014. “International Journal of Information Management Data Quality Management, Data Usage Experience and Acquisition Intention of Big Data Analytics.” *International Journal of Information Management* 34: 387–389.

Letouzé, Emmanuel 2012. “Big Data for Development: Challenges & Opportunities,” *Global Pulse*, no. May.

Li, Guoliang, Beng Chin Ooi, Jianhua Feng, Jianyong Wang, and Lizhu Zhou. 2008. “EASE: An Effective 3-in-1 Keyword Search Method for Unstructured, Semi-Structured and Structured Data.” *Search* pages: 903–914.

Mantelero, A. 2014. “Social Control, Transparency and Participation in the Big Data World.” *Journal of Internet Law*, no. April: 23–29,

Manyika, James, Chui, Michael, Brown, Brad, Bughin, Jacques, Dobbs, Richard, Roxburgh, Charles, and Byers, Angela Hung. 2011. “Big Data: The next Frontier for Innovation, Competition, and Productivity.” *McKinsey Global Institute*, no. June: 156.

Marques, Pedro Miguel Borges. 2016. “Técnicas de Análise de Dados (Data Analytics) No Contexto de Uma Auditoria Financeira.” *OROC*, no. Parte I.

Marr, Bernard, 2013. The awesome ways big data is used today to change our world. <http://www.linkedin.com/today/post/article/20131113065157-64875646-the-awesome-ways-big-data-is-used-today-to-change-our-world> Acesso em 8 de Julho 2016

Martino, Beniamino Di, Rocco Aversa, Giuseppina Cretella, Antonio Esposito, and Joanna Kołodziej. 2014. "Big Data (Lost) in the Cloud." *International Journal of Big Data Intelligence* 1 (1/2): 3.

Mattsson, Ulf T. 2014. "Bridging the Gap Between Access and Security in Big Data." *ISACA Journal*, Vol. 6, pp. 25-29.

Maximodata, Soluções ao problema do *Big Data*. Disponível em: <http://maximodata.com/big-data> Acesso em 28 de Agosto de 2016.

McAfee, Andrew and Brynjolfsson, Erik. 2012. "Big Data. The Management Revolution." *Harvard Business Review* 90 (10): 61–68.

Meadows, Rolling. 2014 "It may be riskier to ignore Big Data than implement it, Says New ISACA White Paper" *ISACA*

Moffitt, Kevin C., and Miklos, A. Vasarhelyi. 2013. "AIS in an Age of Big Data." *Journal of Information Systems* 27 (2): 1–19

Mohapatra, Biswajit, Parisa, Vinay, and Banerjee, Joydipito. 2014. "The Value of Big Data Analytics to the Business." *ISACA Journal*, Vol. 5, pp. 20-26.

Pedrosa, Isabel, and Costa, Carlos J.. 2014. "New Trends on CAATTs: What Are the Chartered Accountants' New Challenges?" *Proceedings of the International Conference on Information Systems and Design of Communication - ISDOC '14*, no. July 2015: 138–142.

Pedrosa, Isabel; Costa, Carlos J., and Raul M S Laureano. 2015. "Use of Information Technology on Statutory Auditors' Work: New Profiles beyond Spreadsheets' Users." *2015 10th Iberian Conference on Information Systems and Technologies, CISTI 2015*, no. June.

Poderoso, Celso. "Visão geral da arquitetura" *Canaltech* black. <http://corporate.canaltech.com.br/coluna/big-data/Visao-geral-da-Arquitetura-de-Big-Data/> Acesso em 26 de Julho de 2016.

Preez, Derek du. 2012. "Maude Introduces Independent Reviews for Open Data Countries." *Computerworld*. <http://www.computerworlduk.com/news/publicsector/3400618/maude-introduces-independent-reviews-for-open-data-countries> Acesso em 25 de Junho de 2016.

Raskin, Victor; Taylor, Julia M., and Christian F. Hempelmann. 2013. "NoMeaning-and Ontology-Based Technologies for High-Precision Language an Information-Processing Computational Systems", *Journal Advanced Engineering Informatics*, Volume 27 Issue 1, January, pp. 4-12.

Reznik, Dan S. 2015 “Big Data, Analytics, Aplicações, Aspectos Práticos e o Cientista de Dados” <https://pt.slideshare.net/dreznik1/big-data-aplicaes-aspectos-prticos-e-cientista-de-dados>

Riffat, Muzamil. 2014. “Big Data—Not a Panacea.” *ISACA Journal*, vol. 3, pp. 19-21.

Russom, Philip. 2013. “Managing Big Data.” TDWI best practices Report - Fourth Quarter 2013, *Tdwi.org*, 3–35.

Russon, Philip. 2011. “Big Data Analytics”, TDWI best practices Report - Fourth Quarter 2011

Sathiyamurthy, Sudhakar. 2011. “A Luta Pela Privacidade E a Segurança Das Informações No Ecosystema de TI.” *ISACA 2*

Schaeffer, Donna M, and Olson, Patrick C. 2014. “Big Data Options For Small And Medium Enterprises.” *Review of Business Information Systems* 18 (1): 41–46.

Schneier, Bruce. 2015. *Data and Goliath The Hidden Battles to Collect Your Data and Control Your World*.

Schonberger, Viktor Mayer, and Cukier, Kenneth. 2013. “Big Data: Como Extrair Volume, Variedade, Velocidade E Valor Da Avalanche de Informação Cotidiana.” In *Big Data: Como Extrair Volume, Variedade, Velocidade E Valor Da Avalanche de Informação Cotidiana*, 203–207.

Schonberger, Viktor Mayer; Cukier, Kenneth. 2013. “Big Data: A Revolution That Will Transform How We Live, Work, and Think.” *New Books in Brief* 7: 242.

Setty, Kumar, and Bakhshi, Rohit. 2013. “What Is Big Data and What Does It Have to Do With IT Audit ?” *ISACA Journal*, vol.5, pp 23-15.

Silde, Alice, and Self, Richard. 2013. “IS Practices for SME Success Series.” *IS Practices for SME Success Series* 1: 1–148.

Singleton, Tommie. 2013. “What Every IT Auditor Should Know About Data Analytics.” *ISACA Journal*, vol 5, pp 12-13.

Singleton, Tommie. 2014. “The Logical Reason for Consideration of IT.” *ISACA Journal*, vol. 3, pp 16-18.

Stonebraker, Michael. 2012. “What Does ‘Big Data’ Mean?” *Communication ACM*.

Stroud, Robert E. 2015. “Research/emerging Technologies: Grab Your Crystal Ball as Disruption Is Certain!,” no. March.

Taurion, Cezar. 2014. *Big Data. Signos Do Consumo*. Vol. 6.

Taurion, Cezar. 2012. “Você Realmente Sabe O Que É Big Data?” <https://www.ibm.com/developerworks/community/blogs/ctaurion/?lang=en>. Acesso a 1 de Julho de 2016.

Titera, William R. 2013. *Updating Audit Standard—Enabling Audit Data Analysis*. Vol. 2.

Trelles, Oswaldo; Prins, Pjotr; Snir, Marc, and Jansen, Ritsert C, 2011. “Big Data, but Are We Ready?” *Nature Reviews Genetics* 12 (3). Nature Publishing Group: 224.

Turner, Vernon; Gantz, John F.; Reinsel, David, and Minton, Stephen. 2014. “The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things.”

Vijayan, Jaikumar. 2012. “Finding the Business Value in Big Data Is a Big Broblem.” *Computerworlduk - The Voice of Business Technology*.

Waller, Matthew A., and Fawcett, Stanley E. 2013. “Click Here for a Data Scientist: Big Data, Predictive Analytics, and Theory Development in the Era of a Maker Movement Supply Chain.” *Journal of Business Logistics* 24: 249–252.

Weigend, Andreas. 2014. “Entrevista Com O Especialista Em Big Data Andreas Weigend: ‘Comece Com as Perguntas.’” *The Huffington Post*.

Yirula, Carolina Prestes, 2013 em <https://cadernodia.wordpress.com/2013/06/18/big-data-uma-possibilidade-para-o-desenvolvimento-social/> Publicado em junho 18, 2013

Yu, William Emmanuel. 2014. “Data Privacy and Big Data—Compliance Issues and Considerations.” *ISACA Journal*, vol. 2, pp. 27-31.

Zimmer, Michael. 2010. “‘But the Data Is Already Public’: On the Ethics of Research in Facebook.” *Ethics and Information Technology* 12 (4): 313–325.

Anexo A. Artigo publicado 11^a CISTI'2016 – Conferência Ibérica de Sistemas e Tecnologia de Informação

Coimbra– Junho 2016

Big Data & Analytics - Requisitos Mínimos de Controlo: uma proposta de revisão da literatura

Big Data & Analytics - minimum control requirements: a proposal on literature review

Vanessa Vieira
Coimbra Business School – IPC
Quinta Agrícola - Bencanta
Coimbra, Portugal
vanessa-vieira@live.com.pt

Isabel Pedrosa
Coimbra Business School – IPC
Instituto Universitário de Lisboa (ISCTE-
IUL), Lisboa, Portugal
ISTAR-IUL
Av. das Forças Armadas, Lisboa
1649-026 Lisboa, Portugal
ipedrosa@iscac.pt

Bruno Horta Soares
Coimbra Business School – IPC
Quinta Agrícola - Bencanta
Coimbra, Portugal
bruno.soares@govaas.com

Resumo — Estamos perante o fenómeno *Big Data* quando nos referimos ao crescimento do volume e variedade de dados, gerados a uma grande velocidade, agregado também à veracidade e valor da informação em função dos avanços tecnológicos. Um dos atuais pontos de interesse no que a este tema diz respeito é o novo desafio no que concerne ao ponto de encontro entre a auditoria e *Big Data*, numa conjuntura de requisitos mínimos de controlo. Este trabalho apresenta conceitos formulados por organizações e especialistas sobre o *Big Data & Analytics*, uma das tendências empresariais, e a sua ligação ao trabalho do auditor, nomeadamente nos desafios que representa para estes profissionais. Os profissionais de auditoria podem vir a beneficiar de grandes transformações com as oportunidades oferecidas pelo *Big Data & Analytics*, permitindo que os recentes avanços de tecnologias de *Analytics* estão a progredir por forma a aprimorar a abordagem e execução de procedimentos.

Palavras Chave – *Big Data*; organizações ; informação; auditoria e tecnologia.

Abstract — We are forwards *Big Data* world, when we refer to the growth and data diversity, generated at high speed, connected to the truth and value of information, all driven by the current technology advances. Another point of interest concerns with the meeting points related with Audit and *Big Data* dimensions, inside a framework of minimum control requirements. This research present precise concepts by organizations and specialists about *Big Data & Analytics*, one of the business trends and their connection to the work of the auditor, including the challenges is for these professionals. Audit professionals can come to benefit from major changes in the opportunities offered by *Big Data & Analytics*, allowing recent advances in *analytics* technologies are progressing in order to improve the approach and implementation procedures.

Keywords – *Big Data*; organizations; information; Audit; technology.

I.Introdução

A nossa sociedade privilegia cada vez mais a informação como uma das suas preocupações mais dominantes[1], apresentando-se como a prioridade para a gestão eficiente de todo e qualquer tipo de organização. Este é um cenário típico dos dias de hoje, pelo que uma das motivações deste trabalho consiste em conhecer de forma mais profunda o cenário da utilização de dados, em particular, no contexto Big Data e perceber o impacto que tal poderá ter na forma como os auditores realizam as suas tarefas. O presente texto introduz o tema e encontra-se dividido em cinco capítulos respeitantes ao assunto:

A primeira secção corresponde à informação introdutória e contextualização dos restantes capítulos.

A segunda secção refere-se à revisão de literatura, introduzindo os conceitos essenciais associados ao Big Data, e assegurando a “ponte” entre a realidade do processamento de dados e o fenómeno Big Data. Nesta secção são clarificados os conceitos, partindo das definições e perspetivas de diversos especialistas na área. Identificam-se, igualmente, os atributos do Big data, ou seja, apresenta, ainda que de forma sucinta, as dimensões que surgem aliadas aos dados digitais e que muitas vezes são designadas por 5 V’s do Big Data. [2]Por fim, esta secção termina com a importância de serem colocadas as perguntas corretas no contexto da Analytics para possibilitar uma correta tomada de decisão.

A terceira secção introduz as áreas fundamentais associadas à Auditoria e controlo interno. Descreve as questões relacionadas com a Auditoria e a Analytics.[3] Mais concretamente, investiga o âmbito de aplicação, realça o vínculo da transformação de dados e a auditoria, e finalmente aborda a Analytics como uma vantagem competitiva para as organizações de auditoria.

Na secção quatro referem-se evidências que fazem parte de um contexto de controlo interno dos dias de hoje[4],

questionando se estarão os grandes dados a moldar a sociedade organizacional.

Por fim, na secção cinco, em modo de conclusão está exposto as considerações finais desta pesquisa. De forma geral, a tecnologia permite analisar qualquer tipo de informação digital em tempo real quer para ajudar nas decisões das organizações, quer para melhorar a competitividade, produtividade e inovação.[5]

II.Revisão de Literatura

A. A realidade do processamento de informação – o fenómeno Big Data

Tão importante quanto gerar informação é a capacidade de processamento de grandes volumes de dados em alta velocidade. Um estudo realizado pela IDC indica que em 2020, o volume de informação digital em circulação, a nível mundial, deverá ultrapassar os 40.000 exabytes. Isto equivale a uma multiplicação por 10 dos valores atuais. [6] Inevitavelmente, o desafio de analisar e entender grandes quantidades de dados está a crescer, criando novos desafios ao armazenamento. O termo Big Data, de forma singela, significa grandes dados, porém, esta expressão encerra desafios significativos para as organizações, sendo capaz de definir o futuro de organizações e organizações no que diz respeito à análise e estruturação de dados.

“Antes acreditávamos que era só começar com os dados, para se chegar aos *insights*, que se traduziriam em conhecimento e eventualmente sabedoria. Hoje sabemos que sabedoria não é suficiente. É preciso começar com as perguntas e então chegar à decisão”, afirma Andreas Weigend, uma das autoridades globais em “Big Data”, 2014 [7]

Na mesma linha, Nate Silver, fundador do FiveThirtyEight, um website de jornalismo de dados, diz que mesmo pequenas quantidades de dados podem ser difíceis de gerir, tanto em termos de armazenamento como de análise. De um modo mais geral, o ser humano, enquanto indivíduo, consome informação

para tomar decisões e, desta forma concretizar em ação as suas intenções. A melhor prática é a verdadeira racionalização dos dados [8].

A. O que é o Big Data

O termo *Big Data* refere-se a um grande conjunto de dados armazenados. Diz-se que o *Big Data* se baseia em 5 Vs: velocidade, volume, variedade, veracidade e valor.

De acordo com Muzamil Riffat: o termo *Big Data* refere-se a um dilúvio de informação que é gerado a cada segundo através de dispositivos digitais que gravam e praticamente ditam os padrões de vida pessoal e profissional de cada um [9]. O real desafio não reside na disponibilidade de dados mas do conhecimento útil a partir deles.

Como visto, o termo é novo porém, para Davenport, "a noção mais ampla que o fundamenta não é particularmente nova" [10]. Para Mayer-Schönberger e Cukier, atualmente "o conceito está a migrar para todos os campos do conhecimento humano" [11] desta forma é necessário conhecer algumas definições (Quadro 1) estabelecidas para o conceito de *Big Data*.

QUADRO 3: DEFINIÇÕES PARA O TERMO *BIG DATA*

<i>Big Data é...</i>	Autor
"(...) Big Dta= Volume + Variedade + Velocidade + Veracidade + Valor"	Taurion (2012)[12]
"As tecnologias de <i>Big Data</i> descrevem uma nova geração de tecnologias e arquiteturas projetadas para extrair economicamente o valor de volumes muito grandes e de uma grande variedade de dados, permitindo alta velocidade de captura, descoberta e/ou análise".	International Data Corporation (2011) [13]
" <i>Big Data</i> refere-se a conjuntos de dados cujo tamanho vai além da capacidade das ferramentas de <i>software</i> de bases de dados típica para capturar, armazenar, gerir e analisar."	Manyika (2011) [5]
" <i>Big Data</i> , em geral, é definido como ativos de alto volume, velocidade e variedade de informação que exigem custo-benefício, formas inovadoras de processamento de informação para maior visibilidade e tomada de decisão."	Gartner Group (2013) [14]
" <i>Big Data</i> é um termo genérico para dados que não podem ser contidos nos repositórios comuns; refere-se a dados volumosos demais para caber num servidor; não estruturados demais para se adequarem a bases de dados organizadas em linhas e colunas; ou fluídos demais para serem armazenados em <i>data warehouses</i> estáticas."	Davenport (2014) [10]

B. Atributos do Big Data

Sistematizam-se agora os atributos que têm sido descritos na literatura como sendo os mais comuns associados a *Big Data*:

- Variedade – Os dados são gerados em todos os tipos de formatos - de dados estruturados, dados numéricos em bases de dados tradicionais, até documentos de texto não estruturados, e-mail, vídeo, áudio, dados de cotações da bolsa e transações financeiras. Ainda podemos classificar os dados em semi-estruturados como, por exemplo, os ficheiros XML [15].

- Volume – caracteriza a quantidade de dados que é gerada continuamente (na ordem de terabyte, petabyte, exabyte, zetabyte). Para Enomura, o volume de dados “já é tão grande que se torna difícil de entender sem o comparar com outros exemplos.” Um estudo realizado em 2014 pela IDC revela como o surgimento das novas tecnologias tem contribuído para o aumento do volume de dados [16].

- Velocidade – caracteriza a velocidade a que esses dados são processados (nomeadamente em tempo real e streams). A rapidez na troca de dados e troca de informação [14].

- Veracidade – a qualidade dos dados e informação é característica essencial para que os utilizadores interessados (na gestão executiva, na gestão pública e na sociedade em geral) usem e (re)usem os dados de forma apropriada e real, gerando informação e, consequentemente, conhecimento. As organizações necessitam de desenvolver melhorias na área de gestão de dados (*Data Governance*) para garantir a veracidade dos dados provenientes das novas fontes [4].

- Valor – Para extrair valor a partir de *Big Data*, a informação necessita de ser verificada para determinar a sua veracidade, de forma que seja possível evitar incerteza. Mayer-Schönberger e Cukier acreditam que o valor real dos dados é como um iceberg que flutua no oceano: apenas parte dele é visível no princípio, enquanto boa parte permanece oculta sob a água. As organizações inovadoras capazes de entender isso extraem o valor oculto – e por isso também, o mais interessante - e criam benefícios significativos [11].

Os atributos anteriormente definidos identificam os principais desafios que precisam de ser considerados quando se considera o termo *Big Data* e a sua análise.

C. *Por detrás de uma grande decisão está sempre uma grande pergunta*

O professor Andreas Weigend, da Universidade de Stanford, afirma que *Big Data* é inútil sem boas perguntas o que nos leva a refletir sobre as questões mais relevantes e úteis que podem ser respondidas com o apoio dessa tecnologia e o nível de maturidade das organizações para elaborar e encontrar respostas a essas perguntas. A visão deste investigador passa por, no futuro, considerar que qualquer decisão tomada por uma organização poderá (e deverá) ser apoiada em dados, números e informação obtida usando técnicas de exploração de *Big Data* [7]

Volume, velocidade, variedade, veracidade e valor são itens que devem ser questionados. Esta combinação, além de todo e qualquer outro aspeto que caracteriza uma solução de *Big Data*, pode ser inviável se o resultado não trazer benefícios significativos e que compensem o investimento das organizações que passam a adotar o *Big Data* como estrutura de Analytics [17].

Antes de armazenar os dados *Big Data* é essencial saber para que ele serve, como funciona, qual a capacidade de armazenamento, se todos os dados da organização precisam de ser armazenados e sobretudo ter conhecimento do serviço oferecido. Assim, se começarmos com uma boa pergunta podemos ser mais assertivos na busca dos dados e consequentemente no resultado do projeto [18].

III.A auditoria e a Analytics

Há uma imensa variedade de necessidades nos negócios de hoje para uma Analytics eficaz. Por esta razão, os auditores têm que ter sensibilidade em determinados fatores que possam

considerar essenciais na condução da Analytics e seguir um programa eficaz análise. [19]

A. Âmbito de aplicação

Segundo Singleton (um investigador com certificações na área de Auditoria de Sistemas de Informação (CISA) e *IT Governance* (CGEIT)), o auditor deve determinar qual a melhor abordagem para satisfazer os objetivos da Analytics, isto é, dar resposta a: como são os dados, onde estão e a melhor forma para os obter tendo sempre em mente a perspetiva de negócio [3].

Este novo conceito de *Big Data* trouxe o estado da arte para a Analytics. As organizações estão inundadas de dados, tanto sob a forma de dados estruturados (ou seja, que estão no seu sistema de gestão) como não estruturados (filmes e emails, por exemplo).[20] Um dos maiores desafios de sempre para as organizações é extrair valor estratégico de toda a imensidão de dados. Assim, quando uma organização implementa uma solução *Big Data* significa que adotou mecanismos para lidar, analisar e mesmo retirar mais-valias do facto de haver tanta informação.[21]

A Analytics pode ser eficaz para auditores na fase de planificação de trabalho de campo [22]. Grande parte do trabalho envolvido na realização de auditorias de TI implica a inspeção de dados gerados a partir de sistemas, dispositivos e outras aplicações.[19] Este é um interessante ponto a ser investigado, principalmente levando-se em conta o atual momento em que aplicações tais como as de inteligência de dados e exploração de ambientes de *Big Data* estão na ordem do dia.

B. A transformação de dados e a auditoria

A análise do perfil e origem dos dados é da competência da Auditoria de Dados. O envio de informação para o exterior da organização é, na maioria dos casos, suportado por relatórios periódicos de operações, relatórios esses que necessitam de auditorias internas prévias ao envio da informação de modo a

garantir que a informação da atividade se encontra em conformidade com os padrões esperados e com a legislação em vigor [23].

Para Kumar Setty, os auditores utilizam ferramentas técnicas de Analytics para tirar conclusões. Ainda para este autor *Big Data* engloba não só as operações realizadas pelas organizações bem como o novo mundo de interações e observações sobre os dados [19]. Adicionalmente, o tipo de utilização destes dados permite analisar de forma rápida e coerente uma grande quantidade de informação em tempo real. [22]. A utilização de *Big Data & Analytics* possibilita agregar mecanismos indispensáveis para a veracidade de dados, ou seja, o controlo interno deve estar sobre a alçada da classificação de fontes de dados em termos de qualidade, precisão e atualização dos dados.

Muitas organizações adotam “grandes análises de dados”, procurando avanços nas técnicas de análise para a rápida expansão do conjunto de informação que as organizações têm à sua disposição para permitir uma melhor tomada de decisão [24]. Torna-se fundamental garantir a segurança da informação. Este trabalho compete aos profissionais de risco e auditores. O risco técnico e operacional surge como resultado da adoção da Analytics por parte das organizações: já a não adoção de Analytics tem influência na competitividade das organizações [25]. Referimo-nos a Analytics considerando um sistema capaz de armazenar, agregar e transformar a ampla gama de fontes de dados da organização. Neste tema, uma funcionalidade potencial é a análise de fraude, as quais representam um risco nas organizações [26].

Novas ferramentas e métodos de Analytics demonstram a possibilidade de as organizações gerarem valor a partir da existência de dados dentro das próprias organizações e de informação externa disponível. Estes avanços permitem melhor tomada de decisão e aumento da vantagem competitiva da organização/organização. No entanto, é fulcral perceber que é um investimento no *Big Data & Analytics* que se destina à

melhoria em termos de eficácia ou eficiência de negócio. Contudo, esta questão envolve risco operacional versus risco de negócio [13].

C. Analytics e vantagem competitiva

A Analytics é fundamental para muitas organizações porque, na maioria das vezes, já foram realizados investimentos na área de análise e recolha de grandes conjuntos de dados, em que a componente analítica pode ser aplicada [5].

Existe uma correlação direta entre o uso de grandes dados e a análise e rentabilidade. Existem estudos que indicam que há aumento de rentabilidade global como resultado do uso de *Big Data* de forma eficaz. A vantagem competitiva está presente caso a organização tenha as pessoas, processos e tecnologias com capacidade para fazer uso dos dados da organização [3]. Para além disso, é fundamental apresentar algum grau de autoconhecimento organizacional é essencial, nomeadamente, pensar e apoiar os processos e identificar potenciais áreas problemáticas antes de serem realizados investimentos significativos em Analytics [4].

Técnicas de Analytics podem conferir vantagem competitiva Empresarial quando se faz jus à avaliação da possibilidade de segurança ao risco de privacidade que poderia ocorrer. Note-se que uma variedade de ferramentas (*software*) são usadas para executar Analytics e ajudam a identificar áreas chave de risco, nomeadamente, fraude, erros de má utilização, permitindo, assim, a melhoria dos negócios, a verificação de eficiência e eficácia dos processos e, por fim, melhoria das decisões associadas ao negócio. Sinteticamente, os benefícios para a organização incluem a eficácia, redução de erros e oportunidade de melhorar a identificação de problemas [6].

A Analytics, quando utilizada de forma eficaz numa organização, está entre os ativos mais valiosos da organização. Extraem-se benefícios se se verificarem estas três premissas: a Analytics está alinhada com o modelo de negócio (*Data Governance*), o risco é gerido e o processo é planeado,

projetado, implementado e testado [15]. No fundo o objetivo é que o conceito *Big Data* converta grandes dados em informação útil. Assim, cumprindo os princípios fundamentais no contexto de *Big Data*, as organizações podem colher os benefícios sem os potenciais riscos para a conformidade no presente e no futuro.

IV. Evidências *Big Data* em controlo interno e auditoria

Um tema atual no contexto da “Era da Informação” é o que se relaciona com a forma como a cada dia (ou mesmo a cada minuto) aumenta a nossa capacidade de produzir dados de todos os tipos [16]. A *Information Week* realizou um estudo envolvendo grandes organizações onde 37,5% afirmaram que a análise de *Big Data* era o seu principal desafio.

As novas possibilidades para a avaliação e resposta a riscos em demonstrações financeiras.[27] Verifica-se que as práticas de *Analytics* permitem que as informações financeiras das entidades sejam analisadas de forma integral e com periodicidades mais curtas, contribuindo para a formulação de opiniões em tempo real e mais confiáveis.[28]

O Controlo Interno e a Gestão do Risco[29] ganharam uma real importância sobretudo diante dos grandes escândalos organizacionais que se tem vivido nesses últimos tempos, onde se procura implementar soluções que passam por *Big Data & Analytics* de modo a mitigar esses impactos.

O controlo interno e Analytics tem sido um fator crucial para as organizações que necessitam de soluções eficientes num mundo globalizado cada vez mais competitivo. [19] *Big Data & Analytics* em auditoria é a estratégia para gerir resultados.

De acordo com a CISCO, 60% dos CIO, Chief Information Office, de organizações a nível mundial concordam que o *Big Data* auxiliará na melhoria da tomada de decisão e aumentará a competitividade no mercado. As aplicações práticas já se encontram para além dos universos mais previsíveis, encontrando-se já em domínios como o desporto: segundo o The

Economist, as equipas da liga inglesa de futebol estão a usar tecnologia *Big Data* para otimizar as suas táticas e a aquisição de jogadores.

V. Conclusões

Com base na investigação realizada neste trabalho, verifica-se que, os avanços tecnológicos, os quais envolvem direta ou indiretamente a sociedade, encontram-se de tal forma avançados que não há como fugir do fenómeno *Big Data*. Torna-se assim um dos principais conceitos nas tecnologias de informação para auditoria sendo já uma realidade muito mais próxima de nós do que possamos imaginar.

A auditoria, os profissionais que controlam o risco de informação e dados, têm um papel fundamental para apoiarem nesta fase em consonância com a utilização apropriada dos dados organizacionais. Assim sendo, o estado de arte é o seguinte: um cruzamento de informação entre o Controlo Interno e *Big Data*, procurando entender os desafios, o âmbito e os riscos que representa para o auditor. Como resultado, o objetivo é auxiliar futuras investigações, baseadas em ferramentas práticas a utilizar por entidades que, eventualmente, possam querer entender quais os objetivos de controlo mais relevantes a implementar / avaliar para responder às atuais necessidades do mercado. Tornando-se, assim, relevante saber como colher frutos do “Fenómeno *Big Data*” e beneficiar do impacto positivo que a *Analytics* pode oferecer. Para a realização deste trabalho foram consultados livros que abordam o assunto *Big Data & Analytics* e artigos de investigação. Apresenta-se uma breve síntese de uma questão que está a vincar de forma relevante, realidade de todo o ambiente organizacional, aspeto que nos leva às dimensões do risco inerente à questão do impacto que causaria a utilização da análise de *Big Data*.

VI. Referências Bibliográficas

- [1] M. Barlow, *Real-Time Big Data Analytics: Emerging Architecture*. 2013.
- [2] SoftServe US Office, “*Big Data Analytics* - Reference Architectures and Case Studies.”
- [3] T. Singleton, “*big data* - the logical reason for consideration of it,” in *big data - the logical reason for consideration of it*.
- [4] U. T. Mattsson, “bridging the gap between access and security in *big data*,” in *bridging the gap between access and security in big data*, 2014, vol. 6, pp. 1–5.
- [5] M. G. Institute, J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers, “*big data* : the next frontier for innovation , competition , and productivity,” 2011.
- [6] V. Turner, J. F. Gantz, D. Reinsel, and S. Minton, “the digital universe of opportunities : rich data and the increasing value of the internet of things,” 2014.
- [7] A. Weigend, “Entrevista com o especialista em *big data* Andreas Weigend : ‘ Comece com as perguntas , ’” 2014.
- [8] L. B. Gouveia and J. Ranito, *sistemas de informação de apoio à gestão*. 2004.
- [9] M. Riffat, “*big data* - not a Panacea,” in *big data - not a Panacea*, 2014, vol. 3, pp. 1–3.
- [10] ISACA, “generating value from *big data analytics*,” in *generating value from big data analytics*, 2014, no. January.
- [11] V. M. Schonberger and K. Cukier, “*big data*: como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana,” in *big data: como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana*, 2013, pp. 203–207.
- [12] “você realmente sabe o que é *Big Data*?,” in *você realmente sabe o que é Big Data?*, 2012, vol. <https://ww>, no. 38691, p. <https://www.ibm.com/developerworks/community/blogs>.
- [13] J. Gantz and D. Reinsel, “extracting value from cgaos,” in *extracting value from cgaos*, 2011, no. June, pp. 1–12.
- [14] B. Mohapatra, V. Parisa, and J. Banerjee, “the value of *big data analytics* to the business,” in *the value of big data analytics to the business*, 2014, vol. 5, pp. 1–7.
- [15] D. M. Schaeffer and P. C. Olson, “*big data* options for small and medium enterprises,” in *big data options for small and medium enterprises*, 2014, vol. 18, no. 1, pp. 41–46.
- [16] S. Kuenkaikaw, “predictive audit *analytics*: evolving to a new era,” in *predictive audit analytics: evolving to a new era*, 2013.
- [17] M. A. Vasarhelyi and K. C. Moffitt, “ais in an age of *big data*,” in *ais in an age of big data*, 2013, vol. 27, no. 2, pp. 1–19.
- [18] B. Franks, “taming the *big data* tidal wave: finding opportunities in huge data streams with advanced *analytics*,” in *taming the big data tidal wave: finding opportunities in huge data streams with advanced analytics*.
- [19] A. O. W. Paper, “*Big Data Analytics* - Advanced *Analytics* in Oracle Database,” in *Big Data Analytics - Advanced Analytics in Oracle Database*, 2013, pp. 1–3.
- [20] Jaikumar Vijayan, “Finding the business value in *big data* is a big problem,” *Comput. - voice Bus. Technol.*, 2012.
- [21] D. du Preez, “Maude introduces independent reviews for open data countries,” *Comput. - voice Bus. Technol.*, no. September 2013, 2012.
- [22] T. Singleton, “what every it auditor should know about data *analytics*,”

what every it Auditor. should know about data Anal., vol. 6, pp. 1–3, 2013.

- [23] T. Singleton, “what every it auditor should know about transforming data for Caats,” in *what every it auditor should know about transforming data for Caats*, 2013, vol. 5, pp. 1–2.
- [24] ISACA, “*big data: impacts & benefits*,” in *big data: impacts & benefits*.
- [25] W. E. Yu, “data privacy and *big data* - compliance issues and considerations,” in *data privacy and big data - compliance issues and considerations*, 2014, pp. 1–5.
- [26] R. E. Stroud, “research/emerging technologies: grab your crystal ball as disruption is certain!,” in *research/emerging technologies: grab your crystal ball as disruption is certain!*, 2015, no. March.
- [27] Gledson Pompeu Corrêa da Costa and T. A. de G. L. Dutra, “Financial Audit the era of *Big Data*: new possibilities for evaluating and responding to risks in financial statements of the Federal Government retransca imagem ctrl + Shift + click para liberar retransca,” pp. 54–61, 2014.
- [28] P. D. Robin Bloor, “*Big Data Analytics* - This time it’s personal.”
- [29] P. Russon, “*Big Data Analytics* - Ad,” 2011.

Anexo B. Entrevista a Auditores

1. Informação do respondente

- 1.1. Função na organização
- 1.2. Número de anos de experiência profissional
- 1.3. Número de anos de experiência no mercado Auditoria

2. Organização onde trabalha

- 2.1. Qual o segmento de mercado/área de atuação que atualmente trabalha/audita?
- 2.2. Neste momento, a organização dispõe de ferramentas para Analytics?

3. Contexto de utilização de *Big Data/Analytics*

- 3.1. A organização tem noção da dimensão dos dados gerados? Ou seja, em média quantos utilizadores têm ativos a gerarem dados diariamente?
- 3.2. Entre 0-50 utilizadores
- 3.3. Entre 50-100 utilizadores
- 3.4. Entre 100-150 utilizadores
- 3.5. Mais de 150 utilizadores
- 3.6. Existe controlo interno constante dessa informação?
- 3.7. Quem na organização faz esse trabalho de recolha, tratamento e análises de dados?
- 3.8. Como e com que ferramentas são extraídas os dados existentes na organização? Qual(s) o(s) *software(s)* utilizado(s)?
- 3.9. A organização utiliza técnicas concretas para extrair, recolher, tratar ou analisar os dados? Se sim quais?
- 3.10. Perante os fatores de risco a ter em conta, como é garantida e gerida a segurança e privacidade dos dados? Por exemplo, se utilizam o armazenamento de dados com proteção para identificação de pessoas, autenticação, etc. de forma a protegerem o acesso à informação
- 3.11. A organização redefine estratégias de negócio com base na estratégia de Analytics?
- 3.12. Dos objetivos da organização, algum está diretamente relacionado com o *Big Data* e/ou Analytics? Se sim quais?

3.13. Está previsto num futuro próximo a contratação de RH para este departamento? Se sim, qual a função e porquê?

4. Utilização do *Big Data* na Auditoria

4.1. Existe orçamento específico para Analytics em Auditoria?

4.2. Quais as razões que levam a organização a investir em *software* de Analytics?

4.3. Que desafios a organização encontra atualmente no tratamento e Analytics? E o que se espera no futuro?

4.4. Como o *Big Data* pode ajudar a auditoria?

4.5. Considera que *Big Data* vai ser uma tendência no futuro utilizado pelos auditores?

4.6. Quais são os principais benefícios que espera do *Big Data* para a sua organização?

4.7. Considera que as tecnologias *Big Data* são eficazes para detetar e investigar riscos? De forma geral, se esta realidade permite um melhor controlo interno no futuro?