



**LISBOA
SCHOOL OF
ECONOMICS &
MANAGEMENT**

MESTRADO
CIÊNCIAS EMPRESARIAIS

TRABALHO FINAL DE MESTRADO
DISSERTAÇÃO

**BIG DATA: IDENTIFICAÇÃO DOS FATORES CRÍTICOS
DE SUCESSO**

JOSÉ PAIVA BATISTA JÚNIOR

JUNHO - 2015



**LISBOA
SCHOOL OF
ECONOMICS &
MANAGEMENT**

**MESTRADO EM
CIÊNCIAS EMPRESARIAIS**

**TRABALHO FINAL DE MESTRADO
DISSERTAÇÃO**

**BIG DATA: IDENTIFICAÇÃO DOS FATORES CRÍTICOS
DE SUCESSO**

JOSÉ PAIVA BATISTA JÚNIOR

ORIENTAÇÃO:

PROFESSORA DOUTORA WINNIE PICOTO (ISEG – UL)

JUNHO - 2015

Agradecimentos

Gostaria de agradecer em primeiro lugar à família, especialmente aos meus pais e irmão, que me apoiou e motivou durante todo o longo processo de realização deste estudo.

À Orientadora, Professora Doutora Winnie Picoto, por toda a sua ajuda, pela paciência e total disponibilidade para esclarecer as minhas dúvidas.

A todos os membros do painel que dedicaram uma parte importante do seu tempo para este estudo.

Por fim, a todos os amigos que sempre me apoiaram ao longo desta jornada.

Resumo

O aumento considerável dos dados disponíveis nas organizações pode ser usado como uma fonte de vantagem competitiva. A tomada de decisão baseada nos dados ganha cada vez mais importância, e cresce também o interesse em torno do *big data*. Apesar da crescente popularidade do conceito do *big data*, a investigação acadêmica realizada nesta temática encontra-se ainda numa fase inicial.

Este estudo pretende complementar a pesquisa existente e tem como objetivo identificar os principais fatores críticos de sucesso para retirar valor do *big data* através da revisão da literatura. Para melhor compreender a relevância de cada um dos fatores identificados, optou-se por colocá-los à disposição de um painel de 18 profissionais de forma a criar uma ordenação de acordo com o grau de importância.

Para este processo foi utilizado a metodologia de rankings Delphi, dividida em três fases e com duas rondas. De acordo com os resultados obtidos os cinco fatores críticos de sucesso mais importantes são: (1) desenvolver uma abordagem de *big data* alinhada com os objetivos estratégicos da organização; (2) possuir infraestrutura de TI adequada; (3) possuir novos modelos e ferramentas capazes de processar grandes volumes de dados; (4) cultura organizacional orientada para a gestão dos dados; (5) análise de dados automatizada ou semi-automatizada.

Palavras-chave: Big Data, Fatores Críticos de Sucesso, Valor Organizacional, Delphi.

Abstract

The increasingly amount of available data in the organization can be used as a source of competitive advantage. Data-driven decision-making gains more and more importance, and the interest around big data keeps growing. Despite this growing popularity of the concept of big data, the academic research conducted on this topic is still at an early stage.

This study seeks to complement the existing research and aims to identify the main critical success factors to extract organizational value from big data. To better understand the relevance of each factor, it was decided to put them at the evaluation of a panel of 18 professionals in order to create a rank accordingly to the degree of importance.

The ranking-type Delphi method was used for this stage of the study. This methodology was divided in three stages and two rounds. Accordingly to the obtained results, the five most important critical success factors are: (1) big data strategy aligned with the business objectives; (2) have the adequate IT infrastructure; (3) have new models and tools capable of processing large amounts of data; (4) Organizational culture oriented towards data management; (5) automated or semi-automated data analysis.

Keywords: Big Data, Critical Success Factors, Organisational Value, Delphi.

Lista de acrónimos

CTO – Chief Technology Officer.

FCS – Fator Crítico de Sucesso.

GPS – Global Positioning System.

IBM – International Business Machines.

IDC – International Data Corporation.

MIT – Massachusetts Institute of Technology.

ROI – Return on investment.

SI – Sistemas de Informação.

TI – Tecnologias de Informação.

Índice

1. Introdução	1
2. Revisão da Literatura.....	3
2.1. Big Data.....	4
2.2. Fatores Críticos de Sucesso em SI	12
2.3. Fatores Críticos de Sucesso para retirar valor do Big Data.....	15
3. Metodologia e Dados	17
3.1. Método Delphi	17
3.2. Primeira Fase.....	20
3.3. Segunda Fase.....	23
3.4. Terceira Fase	25
3.4.1. Terceira Fase – Primeira Ronda	25
3.4.2. Terceira Fase – Segunda Ronda	27
4. Análise e Discussão dos Resultados.....	29
5. Conclusões	34
6. Referências Bibliográficas	36
7. Anexos	42

O presente trabalho foi redigido ao abrigo do novo Acordo Ortográfico.

Índice de Tabelas

Tabela 1 – Lista de FCS para retirar valor do <i>big data</i>	16
Tabela 2 – Valores Kendall W.....	20
Tabela 3 – Composição do Painel	21
Tabela 4 – Tipos de projetos de SI nos quais o painel já esteve envolvido	22
Tabela 5 – Lista dos dez fatores com classificações mais altas	24
Tabela 6 – Ordenação dos fatores após a 1ª ronda	26
Tabela 7 – Ordenação final dos fatores	28
Tabela 8 – Lista completa das classificações dos fatores	42
Tabela 9 – Ordenação alfabética dos fatores	43
Tabela 10 – Estatística descritiva da 3ªfase, 1ªronda	44
Tabela 11 – Resultado do coeficiente W de Kendall 1ªronda	44
Tabela 12 – Estatística descritiva da 3ªfase, 2ªronda	45
Tabela 13 – Resultado do coeficiente W de Kendall 2ªronda	45
Tabela 14 – Resultado do coeficiente de correlação Spearman rho	46

Índice de Figuras

Figura 1 – Processo de estudo Delphi de Schmidt	19
---	----

1. Introdução

As novas tecnologias capazes de recolher e analisar dados permitem conhecer e compreender melhor o cliente, levando a uma tomada de decisão mais informada. De acordo com a IBM (2013a) são gerados diariamente cerca de 2.5 triliões bytes de dados todos os dias. A mesma fonte e a consultora Gartner (2012a) referem ainda que até 2015 serão necessários 4.4 milhões de *data scientists*¹. É neste contexto que se revela a importância de otimizar a gestão dos dados dentro da organização, tornando-se um fator crítico de sucesso na obtenção de vantagem competitiva (Sas, 2011).

A combinação de Internet e democratização da criação e partilha de conteúdos em diversos formatos, originou um grande número de tipos de dados. Os dados aumentaram, desta forma, em volume, variedade e velocidade. A este fenómeno atribuiu-se o nome de *big data* (Baesens et al, 2014).

O potencial impacto do *big data* nas organizações continua a aumentar, ao nível das estruturas e processos de gestão, originando desta forma a necessidade de investigação básica e aplicada entre as diversas disciplinas (tais como por exemplo sistemas de informação, marketing, contabilidade, recursos humanos). Baesens et al (2014) indicam que será necessário desenvolver e melhorar métodos analíticos apropriados para *big data* tendo em conta o volume, variedade, velocidade e veracidade dos dados.

¹ *Data scientist* é um profissional que alia à perspicácia empresarial a capacidade de comunicação das conclusões aos líderes organizacionais assim como aos líderes de TI , de uma forma que pode influenciar a tomada de decisão (IBM, 2014).

Embora muitas organizações estejam a reunir grandes quantidades de dados, são poucas as que conseguem utilizar em pleno os *insights* que estes dados fornecem (Bryant et al, 2008). É ainda necessário ter em conta a crescente importância dos dados para melhorar a oferta de produtos e serviços aos clientes (Buhl et al, 2013). De facto, Heisterberg e Verma (2014) afirmam que “*big data* é a nova fonte de vantagem competitiva em todas as indústrias” (p.242).

A organização Varonis Systems (2012) durante o evento Infosecurity Europe 2012 realizou um inquérito a 180 profissionais do setor de Sistemas de Informação (SI) sobre *big data*. Alguns dos resultados obtidos evidenciam que: 59% dos inquiridos consideravam não existir uma definição clara do *big data* e do seu uso; por outro lado quando questionados se *big data* deveria ser uma prioridade estratégica da organização, 69% respondeu de forma positiva; e 57% indicaram que *big data* seria uma iniciativa estratégica das suas organizações nos próximos cinco anos.

Também a Cisco (2012) realizou um inquérito junto de 1800 profissionais da área de Tecnologias de Informação (TI) para abordar as oportunidades e desafios decorrentes da transformação de dados para vantagens estratégicas. Os resultados indicaram que: 60% dos participantes acreditam que *big data* pode melhorar a tomada de decisão e aumentar a competitividade global; 68% indicaram que projetos de *big data* serão uma prioridade estratégica até, pelo menos, 2018.

Desta forma a questão de pesquisa deste estudo é “quais são os fatores críticos de sucesso mais importantes para retirar valor do *big data*?”. Pretende-

se clarificar os fatores que beneficiam uma estratégia do *big data* assim como os seus benefícios organizacionais, numa altura em que os dados não estruturados crescem 15 vezes mais que dados estruturados (DeRoos et al, 2012) e em que o acesso à informação está democratizado. Ao utilizar a metodologia Delphi, esta dissertação pretende elaborar um estudo exploratório qualitativo nesta temática que a literatura atual reconhece como sendo insuficiente (Agrawal et al, 2012; Baesens et al, 2014).

O estudo apresenta a seguinte estrutura: primeiro, uma revisão da literatura no capítulo dois sobre *big data*, uma breve abordagem sobre fatores críticos de sucesso, fatores críticos de sucesso de SI, e por fim um foco específico nos fatores críticos de sucesso para retirar valor do *big data*. O capítulo três aborda a metodologia, a apresentação do painel, assim como a recolha e tratamento de dados. O capítulo quatro apresenta a análise dos resultados obtidos, discussão e confrontação com a literatura. Por fim o capítulo cinco expressa as conclusões, implicações práticas, as limitações e pesquisa futura associada à temática em estudo.

2. Revisão da Literatura

A revisão da literatura conta com três secções. A primeira irá centrar-se em *big data* onde são abordadas as suas definições e características, a sua origem e o seu impacto organizacional. Uma vez que este estudo pretende abordar os fatores críticos de sucesso para retirar valor do *big data* é definido o conceito de fator crítico de sucesso, sendo esta a segunda secção. Por outro lado é importante também analisar o que a literatura existente refere sobre os

fatores críticos de sucesso associados ao setor dos Sistemas de Informação. Por fim na última secção serão então apresentados um conjunto de fatores críticos de sucesso para retirar valor do *big data* com base na análise da literatura existente.

2.1. Big Data

Durante a década de 1980 dá-se uma alteração do paradigma de sistemas centralizados para sistemas distribuídos. Esta mudança fomentou o aumento dos dados distribuídos (Ferguson, 2013). Mais tarde, dois grandes desenvolvimentos levaram ao surgimento do conceito *big data*. Em primeiro, a quantidade de dados gerados aumentou consideravelmente. Em segundo lugar, o hardware conseguiu aumentar a sua capacidade ao mesmo tempo que diminui o preço, tudo isto seguindo a comoditização das tecnologias de informação (Verheij, 2013).

As principais fontes do *big data* geram novos tipos de dados não estruturados ou semi-estruturados. Estima-se mesmo que 80% dos dados mundiais são não estruturados, o que torna a análise e recolha de *insights* um enorme desafio. O súbito aumento na utilização de redes sociais, dados gerados por máquinas, e dispositivos *smart* adicionaram mais complexidades na gestão do *big data* e tornam mais complexa a obtenção de valor (Heisterberg e Verma, 2014). A quantidade de dados nas organizações aumenta, também porque são atualmente capturados de uma forma mais detalhada (Fawcett e Waller, 2013).

De acordo com Lohr (2012) grandes quantidades de dados irão passar pela academia, pelos negócios e pelos governos. Serão poucas as áreas a

permanecerem intocadas pela aumento destes fluxos de dados. O autor refere que de acordo com a IDC, anualmente a quantidade de dados aumenta cerca de 50%.

Os avanços nos sensores digitais, comunicações, computação e armazenamento de dados originaram enormes conjuntos de dados, capturando informação que pode ser valiosa para o negócio das organizações, para a ciência, para o governo e para a sociedade. Por exemplo, motores de pesquisa como Google, Yahoo e Bing, criaram novos modelos de negócios ao capturar informações gratuitas disponíveis na *World Wide Web* (Bryant et al, 2008). É necessário ter ainda em conta que muitos destes dados envolvem aspetos do comportamento do cliente difíceis de observar anteriormente (Meyer et al, 2014).

Nos próximos cinco anos deverá ocorrer um aumento na adoção do *big data*. O volume de negócios do segmento deve aumentar cerca de 400%, traduzindo-se em cerca de 16 mil milhões de euros até 2016 (Buhl et al, 2013). Cada vez mais as organizações irão preocupar-se em formular estratégias para *big data* (Fattah, 2014). Este ambiente de rápidos desenvolvimentos tecnológicos levanta sérias questões aos executivos sobre como ajudar a organização a capitalizar com as transformações em curso (Bughin et al, 2010).

O termo *big data* surge na década de 1970 sendo utilizado para descrever grandes quantidades de dados gerados por experimentações meteorológicas (Heisterberg e Verma, 2014). No entanto, é atribuída à consultora Gartner a popularização do termo em 2001 (Verheij, 2013). Este termo definia os desafios e oportunidades no aumento dos dados, tendo em

conta o aumento de volume, velocidade e variedade. Em 2012 a mesma consultora atualizou esta definição, afirmando que *big data* são ativos de informação com grande volume, velocidade e/ou variedade, e que necessitam de novas formas de processamento para melhorar a tomada de decisão, a descoberta de *insights* e otimização de processos (Verheij, 2013).

Heisterberg e Verma (2014) classificam *big data* como um novo paradigma tecnológico para criar agilidade nos negócios e previsões, ao analisar dados originários de várias fontes. *Big data* para Lohr (2012) pode ser entendido como um termo de marketing mas ao mesmo tempo uma abreviação de tendências tecnológicas que abrem a porta a novas abordagens para compreender informações e tomar decisões. Para Buhl et al (2013) *big data* é uma fusão multidisciplinar e evolutiva de novas tecnologias em combinação com novas dimensões de armazenamento e processamento de dados, uma nova era de variedade de dados e por fim os desafios de gerir a qualidade dos dados. Kandavel e Sundaresan (2014) afirmam que *big data* são dados que excedem a capacidade de gestão de um sistema convencional de base de dados, ou seja dados para quais a arquitetura tradicional não se adequa.

Boyd e Crawford (2012) definem *big data* como um fenómeno tecnológico, cultural e académico, que está baseado na interação de: (1) tecnologia: deve maximizar o poder computacional e precisão dos algoritmos para recolher, analisar, ligar e comparar conjuntos de dados; (2) análise: analisar padrões para melhorar a tomada decisão; (3) “mitologia”: crença que grandes conjuntos de dados oferecem uma forma superior de *intelligence* e conhecimento que geram *insights* outrora impossíveis.

Brown et al (2011a) alertam para o facto da definição do *big data* ser moldada consoante cada setor de atividade, devido às ferramentas habitualmente disponíveis e o tamanho dos conjuntos de dados variarem de acordo com cada indústria. Sendo assim, *big data* em muitos setores varia entre dezenas de terabytes a múltiplos petabytes.

Barton et al (2012) evidenciam que existe alguma confusão entre os conceitos de *big data* e de *analytics*. Para estes autores *big data* diferencia-se tendo em conta três aspetos chave: (1) volume: a partir de 2012, são criados cerca de 2.5 exabytes de dados todos os dias. Isto mostra como as organizações devem trabalhar com grandes quantidades de dados. (2) Velocidade: a informação em tempo real ou muito próxima disso permite à organização tornar-se mais ágil que a concorrência. (3) Variedade: *big data* ganha diversas formas como mensagens, atualizações ou imagens em redes sociais, leituras de sensores, sinais de GPS, entre outras. De facto algumas fontes do *big data* são relativamente recentes. Como é o caso das redes sociais, *smartphones*, outros aparelhos móveis, que são capazes de fornecer um grande fluxo de dados. Brown et al (2011b) também afirmam que atualmente os dados provém de múltiplos pontos e dispositivos presentes na cadeia de valor.

A literatura indica ainda mais duas dimensões do *big data*: (4) veracidade: é difícil assegurar a boa qualidade dos dados (IBM, 2013b; Lajara et al, 2013), e (5) valor: existe uma grande variação no valor de diferentes dados. Os investimentos em *big data* podem ainda oferecer mais valor ao possibilitar o acesso a grandes quantidades de dados em tempo real

(Heisterberg e Verma, 2014). Este valor pode ser alcançado, por exemplo, ao auxiliar a tomada de decisão que melhora o desempenho financeiro, organizacional ou operacional.

No entanto *big data* apresenta ainda desafios que não foram resolvidos relativos ao volume, velocidade e veracidade dos dados (Buhl et al, 2013; Gartner, 2012b).

A literatura ainda não consegue identificar todo o valor possível de retirar do *big data*, e quantificar os seus benefícios poderá ser ainda mais difícil. No entanto, a título exemplificativo, de acordo com Brown et al (2011a) os retalhistas capazes de fazer uso de todo o potencial do *big data*, têm a possibilidade de aumentar a sua margem operacional em cerca de 60%. A mesma fonte indica ainda que *big data* tem o valor anual aproximado de 250 mil milhões de euros para o setor da administração pública da Europa. Ao abordar o *big data*, a organização pode extrair valor e *insights* valiosos de dados volumosos para além do que era possível anteriormente (Heisterberg e Verma, 2014).

Lohr (2012) identifica as organizações Google e Facebook como os atuais líderes na potencialização de dados provenientes da internet. Por sua vez a capacidade preditiva do *big data* está a ser explorada essencialmente pelos campos da saúde, e do desenvolvimento económico e previsões económicas. O autor cita ainda o Professor Brynjolfsson quando este refere que na esfera empresarial, económica e em diversos outros campos, a intuição e experiência irão dar lugar aos dados e sua análise na tomada de decisão.

Barton et al (2012) afirmam que devido ao *big data* os gestores sabem

mais sobre os seus negócios, traduzindo-se no melhoramento da tomada de decisão e desempenho organizacional. Um exemplo prático da melhoria da tomada de decisão e desempenho organizacional pode ser visto no retalho. Uma livraria antes só conseguia acompanhar as vendas e clientes através de programas de fidelização. Com a mudança para plataformas online tornou-se possível conhecer mais o cliente: visualizar não só aquilo que comprou mas também outros artigos que pesquisou, como navegou pelo site, como reagiu a promoções entre outros aspetos. Tornou-se, assim, possível desenvolver algoritmos capazes de sugerir aos clientes itens que vão ao encontro dos seus hábitos de pesquisa. Desta forma, para estes autores, *big data* tem também a capacidade de transformar pequenas e médias organizações.

Meyer et al (2014) citam uma entrevista de Einav (2013) quando este foi questionado sobre a reação das organizações relativamente ao *big data* afirmou que: “(...) muitas delas apenas “sentam-se” em cima dos seus conjuntos de dados e apercebem-se que têm provavelmente uma mina de ouro de dados, mas não têm qualquer ideia do que fazer com eles” (p. 4).

Para avaliar os possíveis impactos do *big data* nas organizações, Brown et al (2011b) analisaram 20 setores da economia norte americana. Os autores concluíram que existem setores com maior potencial para criar valor a partir do *big data*. Os *players* financeiros são os que apresentam mais oportunidades de criar valor a partir do *big data*. No extremo oposto encontram-se os serviços educativos e artísticos. Os autores preveem ainda que *big data* pode melhorar os índices de produtividade organizacionais entre 0,5% a 1% anuais.

A Ford Motor, a Pepsi e a Southwest Airlines são outros exemplos de

adaptação aos novos fluxos de dados. Estas organizações analisam os *posts* acerca delas nas redes sociais para medir o impacto imediato das suas campanhas de marketing e para compreender como o sentimento do consumidor se altera em relação às suas marcas (Bughin et al, 2010).

A consultora Gartner (2012b) identificou quatro principais impactos do *big data* nas organizações: (1) permite que a tomada de decisão identifique padrões de forma rápida em diversos tipos de dados; (2) expõe silos internos que a gestão de topo deve abordar através de incentivos e métricas que encorajam a partilha e confiança nos dados; (3) expõe *gaps* de talentos, introduz novos papéis interdisciplinares, e força a organização a atrair e reter especialistas e gestores de dados com fortes capacidades analíticas; (4) vai obrigar os especialistas em tecnologias a preocuparem-se também com a aplicação de ferramentas, técnicas e arquiteturas para analisar, visualizar, ligar e gerir grandes quantidades de dados.

Há muita informação por extrair do *big data* para potenciar o crescimento da organização (Heisterberg e Verma, 2014). Brown et al (2011b) afirmam que *big data* pode também modificar o ambiente competitivo essencialmente devido a transformar processos de negócio, alterar ecossistemas organizacionais, e facilitar inovações. Os autores concluem que *big data* pode de facto tornar-se um novo tipo de ativo da organização.

No entanto é necessário ter em conta os aspetos menos positivos decorrentes do *big data*. O risco de falsas descobertas, tendo em conta a dificuldade de encontrar e entregar valor em grandes quantidades de dados, aumenta consideravelmente (Lohr, 2012). Por outro lado, os desafios para

tornar a organização orientada para *big data* podem ser enormes (Barton et al, 2012). Não devendo ser esquecido também os recursos financeiros necessários para abordar *big data*.

Uma vez que o acesso à informação tem vindo a ser melhorado, este é um fator incentivador da tendência da crescente utilização do *big data*. Para Lohr (2012) os dados estão não só a tornarem-se mais acessíveis como mais compreensíveis para os computadores. O grande aumento de dados do *big data* são na sua maioria dados não estruturados. No entanto, ao mesmo tempo aumentam as ferramentas informáticas para recolher informações e *insights*. Alguns exemplos são as técnicas de inteligência artificial como processamento de linguagem natural, reconhecimento de padrões e aprendizagem automática. Estas são tecnologias de inteligência artificial possíveis de aplicar a diversos campos.

Por outro lado as soluções de armazenamento de dados também estão a tornar-se mais barata e comuns (Barton et al, 2012). Os investimentos em ferramentas específicas para *big data* devem ocupar cerca de 11% dos gastos em TI nas organizações em 2020 (Lajara et al, 2013).

Na sua tese de doutoramento Verheij (2013) realizou oito estudos de caso em organizações holandesas, e identificou cinco fases distintas para a adoção de soluções *big data*: (1) fase de desenvolvimento estratégico; (2) fase de desenvolvimento do conhecimento; (3) fase de teste piloto; (4) fase de implementação da plataforma; (5) fase de aperfeiçoamento.

Barton et al (2012) concluem que *big data* pode ser de facto o início de um novo paradigma organizacional ao afirmarem que:

à medida que as ferramentas e filosofias do big data se alastram, elas vão alterar ideias de longa data sobre o valor da experiência, a natureza da perícia, e práticas de gestão. Líderes inteligentes em todas as indústrias serão vistos a utilizar o big data para o que é: uma revolução na gestão. (p. 1)

Para este estudo optou-se por adotar a seguinte interpretação do *big data* apresentada por Brinkkemper et al (2013): dados caracterizados por alto volume, alta velocidade e grande variedade, que por conseguinte requerem ferramentas e técnicas exclusivas para gerir dados, processar, armazenar, analisar e visualizar. Optou-se ainda por considerar a veracidade (IBM, 2013b) como a quarta característica essencial do *big data*.

2.2. Fatores Críticos de Sucesso em SI

O conceito de fatores críticos de sucesso adquiriu diferentes interpretações desde que ficou popularizado na década de 1980 através do autor Rockart (Ahmed, 2014). Este autor defini-os como áreas cujos resultados, caso sejam satisfatórios, permitem um desempenho competitivo da organização bem sucedido. Desta forma é essencial identificar quais são os fatores críticos. Os fatores críticos de sucesso podem ainda ser utilizados como um auxílio nos processos de planeamento da organização, para melhorar a comunicação (interna ou externa), e para ajudar no desenvolvimento dos sistemas de informação (Bullen e Rockart, 1981). Para Belassi e Tukul (1996) existe um conjunto de fatores fora do controlo da gestão, e que podem determinar o sucesso ou não de um projeto. São assim os fatores críticos de sucesso ou insucesso. Lynch (2003), por sua vez, vê os fatores críticos de sucesso como recursos, capacidades, e atributos organizacionais que são

essenciais para o desempenho. Pal e Torstensson (2011) indicam que as organizações desenvolvem fatores críticos de sucesso diferentes, de acordo com a estrutura, estratégia competitiva, posição na indústria, localização, e fatores ambientais e temporais, integrados com os objetivos estratégicos. É já reconhecida a importância dos fatores críticos de sucesso, uma vez que eles acabam por influenciar a realização dos objetivos estratégicos da organização (Bullen e Rockart, 1981).

Deng et al (2013) indicam que cerca de 44% dos projetos de SI são entregues fora do prazo, acima do orçamento e sem cumprir os requisitos. Adicionalmente, cerca de 24% são cancelados antes da conclusão ou são entregues mas nunca utilizados. Estes dados evidenciam a necessidade de ultrapassar uma multiplicidade de obstáculos para atingir resultados bem sucedidos.

Rockart (1982) define os fatores críticos de sucesso em SI como pequenas declarações das áreas de número limitado onde “tudo deve correr bem” para que a função de SI tenha sucesso. O autor realizou estudos de caso junto de nove organizações, e após as entrevistas o autor chegou a duas principais conclusões: em primeiro, os fatores críticos de sucesso de SI diferem de organização para organização, no entanto é possível identificar um conjunto comum de quatro fatores críticos de sucesso. Em segundo, cada departamento de SI estabelece as suas ferramentas de gestão, técnicas e processos para alcançar objetivos nas áreas críticas.

O conjunto de quatro fatores críticos de sucesso de SI identificados em todas as organizações do estudo foram: (1) as ferramentas, técnicas e processos

devem estar nos locais certos da organização para que a gestão das pessoas de SI seja eficaz; (2) deve-se garantir que as pessoas se encontram a trabalhar no projeto correto. Com recursos limitados, as prioridades da organização devem ser selecionadas com atenção; (3) os serviços entregues pelos SI devem ser a tempo e dentro do orçamento. Assim como o utilizador deve perceber que o serviço que lhe é entregue é bom, caso contrário a qualidade do serviço torna-se irrelevante; (4) por fim as prioridades de SI devem estar alinhadas com as necessidades de negócios e objetivos organizacionais.

Por sua vez, o conceito de sucesso em SI pode ser entendido como a eficiência final na realização da tarefa para a qual o sistema de informação é desenvolvido, de acordo com Caldeira e Ward (2003). Os autores realizaram um conjunto de estudos de caso a 12 pequenas e médias empresas portuguesas sobre o sucesso na adoção de SI. Dos resultados os autores identificaram dois conjuntos de fatores: (1) fatores secundários, que são importantes mas não críticos para atingir sucesso. Exemplos destes fatores são disponibilidade de recursos financeiros, qualidade do *software* disponível no mercado, disponibilidade e qualidade dos recursos humanos; (2) o segundo conjunto de fatores foram classificados como determinantes, uma vez que evidenciam claramente o porquê de certas organizações terem mais sucesso na adoção de SI. Estes fatores são as competências de SI (pessoas e conhecimento disponível em SI), e as atitudes e perspectivas da gestão face à adoção e utilização de SI. As organizações com menos sucesso na adoção de SI colocaram um ênfase maior na qualidade percebida do *software* que

adquiriam e não no desenvolvimento de competências de SI dentro da organização.

Turner (2014) indica que uma abordagem pessoas, processos e tecnologias maximiza as hipóteses de sucesso na implementação de SI, ao mesmo tempo que potencia os benefícios organizacionais esperados pela gestão de topo.

2.3. Fatores Críticos de Sucesso para retirar valor do Big Data

Para este estudo adotou-se interpretação de fatores críticos de sucesso como os fatores essenciais para atingir sucesso em determinada atividade (Rockart, 1982). Enquanto que retirar valor entenda-se obter *insights*, informações e originar conhecimento relevante para a organização (e.g. extrair um padrão de consumo; identificar uma necessidade que se traduza num produto novo; identificar uma ameaça). Desta forma pretende-se identificar quais os fatores críticos de sucesso mais importantes para que seja possível à organização retirar valor do *big data*. Para esta componente do estudo foram analisados artigos científicos, livros, conteúdos de diferentes tipo, entre eles *press releases*, artigos online e notícias.

Tendo em conta o interesse atual em *big data* pouco foi escrito até agora sobre fatores críticos de sucesso para obter valor organizacional. Desta forma assume-se relevante elaborar um estudo sobre esta temática. Estes fatores foram identificados de acordo com a opinião dos respectivos autores como muito importantes e essenciais a obtenção de valor. Devido à extensão do conjunto de fatores críticos de sucesso (FCS) identificados, a seguinte tabela apresenta uma síntese desta fase da investigação:

1. Análise de dados automatizada ou semi-automatizada.	Bryant et al (2008).
2. Capacidade de inovação e adaptação do modelo de negócio.	Brown et al (2011b); Buhl et al (2013).
3. Cultura organizacional orientada para a gestão de dados.	Barton et al (2012); Buhl et al (2013).
4. Desenvolver uma abordagem de <i>big data</i> alinhada com objetivos estratégicos da organização.	Rijmenam (2013a).
5. <i>Enterprise Architecture</i> estar preparada para <i>big data</i> .	Gartner (2012b).
6. Evitar acumulação de dados em silos departamentais.	Brown et al (2011a).
7. Existir alta qualidade dos dados.	Bughin et al (2010); Buhl et al (2013); Heisterberg e Verma (2014); Verheij (2013).
8. Existir recursos humanos com capacidades para abordar <i>big data</i> .	Brown et al (2011a); Brown et al (2011b); Fattah (2014); Fawcett e Waller (2013); Verheij (2013).
9. Facilitar acesso dos <i>stakeholders</i> aos dados.	Brown et al (2011b).
10. Obter <i>sponsorship</i> para projetos <i>big data</i> .	Rijmenam (2013b).
11. Organização deve permitir e incentivar a experimentação.	Brown et al (2011a); Brown et al (2011b); Bughin et al (2010).
12. Organização deve ser capaz de implementar soluções <i>big data</i> escaláveis.	Verheij (2013).
13. Organização deve ter plataforma de experiência do cliente multi-canal.	Heisterberg e Verma (2014).
14. Possuir infraestrutura de TI adequada.	Brown et al (2011b); Bryant et al (2008); Buhl et al (2013); Heisterberg e Verma (2014); Verheij (2013).
15. Possuir novos modelos e ferramentas capazes de processar grandes volumes de dados.	Buhl et al (2013); Heisterberg e Verma (2014).
16. Possuir práticas de <i>Business Intelligence</i> (tais como análise preditiva e <i>big data analytics</i>).	Brinkkemper et al (2013); Fawcett e Waller (2013); Heisterberg e Verma (2014); Lycett (2013); Verheij (2013).
17. Realizar investimentos e políticas organizacionais orientadas para o futuro.	Brown et al (2011a).
18. Reduzir custos da gestão de dados mas aumento da velocidade de tratamento.	Buhl et al (2013).
19. Respeitar segurança e privacidade dos dados.	Accenture (2011); Bryant et al (2008); Buhl et al (2013); Verheij (2013).
20. Tomada de decisão suportada em algoritmos automatizados.	Brown et al (2011b).

Tabela 1 – Lista de FCS para retirar valor do *big data*.

3. Metodologia e Dados

A principal questão de pesquisa deste estudo é “quais são os fatores críticos de sucesso mais importantes para retirar valor do *big data*?”. Seguindo a recolha de fatores críticos de sucesso baseada na revisão da literatura, esta pesquisa propõe um estudo exploratório qualitativo e uma abordagem metodológica de rankings Delphi (Schmidt, 1997), para melhor compreender o tema em estudo.

3.1. Método Delphi

Para este estudo considerou-se a metodologia Delphi de rankings que usa estatística não paramétrica e segue a linha orientadora proposta por Schmidt (1997) como a mais apropriada por diversas razões. Em primeiro lugar porque será necessário expor os pressupostos e informações que levam a interpretações diferentes e procurar gerar consenso num grupo (Hasson et al, 2000). Em segundo lugar a metodologia Delphi é uma forma apropriada para trabalhar com opiniões e perspetivas ao invés de factos objetivos (Schmidt, 1997). Em terceiro lugar, adequa-se a estudos que procuram investigar uma questão incerta e ou especulativa, onde a população geral não é capaz de responder às questões levantadas pela pesquisa (Okoli e Pawlowski, 2004). Por fim, porque pretende-se apresentar os resultados de uma forma que a sua compreensão seja fácil, tornando-os acessíveis a investigadores, académicos e atores do setor (Schmidt, 1997). Os rankings Delphi podem ainda ser vantajosos para a fase inicial de desenvolvimentos teóricos, como é o caso do

big data, uma vez que auxiliam o investigador a identificar as variáveis de interesse e gerar proposições (Okoli e Pawloski, 2004).

A metodologia Delphi surge na década de 1950 quando a RAND Corporation procurava uma técnica para obter consenso de um grupo de especialistas de forma fiável (Okoli e Pawloski, 2004).

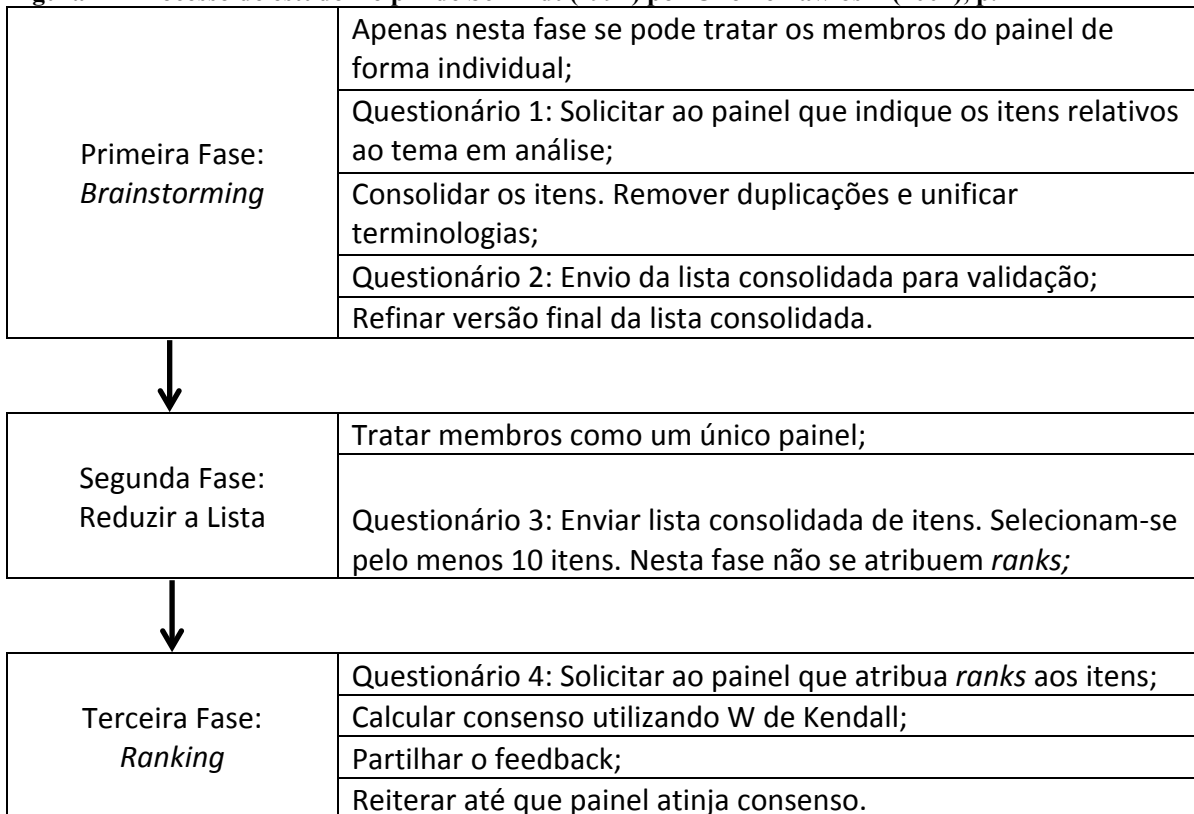
O método Delphi é um processo de tomada de decisão em grupo, a partir do feedback iterativo, onde as opiniões de um painel de peritos são evidenciadas e é atingido um consenso (Deng et al, 2013).

A tipologia de rankings do método Delphi não possui um esquema de trabalho específico, no entanto Schmidt (1997) propõe a divisão desta metodologia em três fases distintas (figura 1). A primeira fase centra-se na recolha de dados e escolha dos membros do painel. A segunda fase organiza a lista de questões para que seja possível criar um ranking. Por fim na terceira fase é apresentada a lista, de forma aleatória, e solicitado ao painel que ordene os itens tendo em conta o nível de importância.

Para a utilização desta metodologia é necessário ter em conta três aspetos importantes (Schmidt, 1997): em primeiro lugar o investigador deve perceber quando deve parar de realizar rondas. Se parar demasiado cedo, o ranking pode não ter valor mas demasiadas rondas podem desmotivar o painel de membros e prejudicar os resultados. De acordo com Hasson et al (2000) caso o número de respostas diminua à medida que as rondas avançam, deve ser ponderada a suspensão de rondas futuras. Em segundo lugar está a quantidade de itens a transportar para a terceira fase. Demasiados itens podem

dificultar a tarefa de atingir consenso. Em terceiro o investigador deve fazer bom uso de técnicas estatísticas para suportar as suas conclusões.

Figura 1 – Processo de estudo Delphi de Schmidt (1997) por Okoli e Pawloski (2004), p.24



Para medir o consenso utiliza-se o coeficiente de concordância de Kendall *W*. Este coeficiente permite determinar de forma realista se foi atingido consenso; se o consenso aumentou ou diminuiu entre rondas; e a força relativa do consenso. Este método é o mais popular em grande parte devido à sua simplicidade de aplicação. Schmidt (1997) propõe a seguinte interpretação dos valores de Kendall *W*:

Tabela 2 – Valores Kendall W. Schmidt (1997), p.767

W	Interpretação	Confiança no ranking
0,1	Concordância muito fraca	Nenhuma
0,3	Concordância fraca	Baixa
0,5	Concordância moderada	Razoável
0,7	Concordância forte	Alta
0,9	Concordância forte fora do normal	Muito alta

Após cada ronda é necessário ponderar se uma nova ronda pode reunir maior consenso. É essencial, no entanto, ter em conta a viabilidade e ganho potencial de uma nova ronda. Quando o W de Kendall é pequeno esta decisão é mais fácil de tomar. No entanto quando há consenso moderado torna-se mais difícil pesar as vantagens e desvantagens.

Para avaliar a correlação de opiniões entre rondas será utilizado o coeficiente de correlação da ordem de ranking de Spearman ρ , uma vez que este enfatiza a magnitude de diferença entre as posições dos itens no ranking (Schmidt, 1997). Os valores deste teste variam entre -1 e 1, sendo que quanto mais próximo do valor 1, maior será a convergência das opiniões entre rondas. Por outro lado o valor -1, indica que não existe entendimento nas ordenações entre rondas (Kalaian e Kasim, 2012).

3.2. Primeira Fase

A primeira fase centra-se na recolha de informação sobre os membros do painel, e a seleção e compilação dos itens da lista (Deng et al, 2013).

O tamanho ideal do painel é dez a dezoito pessoas e no máximo quatro painéis ao mesmo tempo (Okoli e Pawloski, 2004). A constituição do painel teve como objetivo a escolha de profissionais conhecedores da temática em estudo, e baseou-se ainda em Okoli & Pawloski (2004) quando afirmam que o

mesmo deve ser formado principalmente por pessoas com qualificações para responder aos inquéritos elaborados pelo investigador. A identificação destes profissionais foi feita através de contactos diretos com consultoras e profissionais do setor de SI, e responsáveis académicos. O painel foi composto por: seis profissionais alocados a projetos e atividades do *big data*, cinco consultores de *business intelligence*, três docentes de SI, um jornalista especializado em SI, um CTO de uma organização especializada em consultoria TI, um *partner* de uma organização focada em desenvolvimento e implementação de soluções de software, e um diretor de uma *start up* portuguesa que opera também no setor TI. Relativamente à caracterização do painel de acordo com as habilitações literárias verificou-se que 39% possui licenciatura, 28% mestrado, 22% doutoramento e 11% pós-graduação.

Tabela 3 – Composição do Painel

Variável	Perfil do Painel
Demográfica	
Género	
Feminino	4
Masculino	14
Média anos de experiência no setor SI/TI	
1-5 anos	6
6-10 anos	6
11-15 anos	1
16-20 anos	2
Mais de 20 anos	3

Por sua vez a caracterização do painel de acordo com os tipos de

projetos de SI nos quais já estiveram envolvidos, apresentado pela seguinte tabela, evidencia a experiência e relevância dos constituintes para o estudo:

Tabela 4 – Tipos de projetos de SI nos quais o painel já esteve envolvido.

Tipos de Projetos SI	Painel
Análise de Sistemas de Informação	72%
Bases de Dados	67%
Análise de Dados	61%
Desenvolvimento de Software	50%
Sistemas de Apoio à Decisão	50%
Gestão de Projetos Informáticos	50%
Integração de Sistemas	50%
Gestão de Sistemas de Informação	44%
Manutenção	44%
Planeamento e Arquitetura	39%
Gestão de Relacionamento com o Cliente	39%
Rede	22%
Infraestrutura	22%
<i>E-Business</i>	17%
<i>IT Procurement</i>	11%

Todos os membros do painel foram contactados antecipadamente via email, onde lhes foi explicada a metodologia e os objetivos, de forma a garantir a participação neste estudo.

Nesta fase Schmidt (1997) propõe que seja solicitado ao painel de forma aberta, a identificação das principais questões a serem tratadas nas rondas futuras. No entanto para o presente estudo optou-se por identificar as principais questões a partir da revisão da literatura. O resultado foi a lista apresentada na Tabela 1, presente na revisão da literatura.

Optou-se por criar a lista com 20 itens uma vez que pareceu ser uma quantidade satisfatória, de acordo com três aspetos importantes dos rankings Delphi identificados por Hasson et al (2000): deve ser tido em conta o tempo

disponível para o estudo; a complexidade e quantidade de questões que o painel deve responder e por conseguinte a fadiga do painel. Chegou-se assim à conclusão que mais itens poderiam desmotivar a participação neste estudo.

3.3. Segunda Fase

Schmidt (1997) indica que nesta fase deve-se atingir uma lista com 10 a 20 itens no máximo. Sendo assim optou-se por estabelecer como objetivo desta fase identificar os 10 fatores críticos de sucesso para retirar valor do *big data* mais importantes de acordo com o painel. Foi apresentada a lista com os 20 fatores identificados e solicitado que o painel classificasse cada um de acordo com uma escala numérica tipo Likert (Likert, 1932; Boone e Boone, 2012). Os valores apresentados pela escala iam de um, equivalente a nada importante, até dez, equivalente a muito importante. A literatura apresenta diversas visões diferentes relativas ao tamanho ideal das escalas. Para Munshi (2014) o tamanho ideal da escala deve refletir acima de tudo qualidade, ou seja a capacidade de demonstrar fielmente a atitude/opinião que se encontra a ser medida. Desta forma é necessário procurar um equilíbrio: escalas com poucas classificações podem perder as capacidades discriminativas dos inquiridos; por outro lado escalas com demasiadas classificações podem ir para além das verdadeiras capacidades dos respondentes (Jacoby e Matell, 1971). Como foi referido, para este estudo optou-se por escalas com dez classificações. De acordo com Bayer e Wittink (1994) as escalas com dez pontos apresentam as seguintes vantagens quando comparadas com as escalas mais utilizadas de cinco e sete pontos: oferecem variações maiores; maior precisão de medição; fornecem melhores oportunidades de detetar mudanças e conferem mais

poderes para explicações. As escalas maiores também ajudam a minimizar os problemas de benevolência, tendência central e o efeito “halo” (Walker, 1995).

Tendo em conta a disponibilidade horária dos membros do painel decidiu-se utilizar inquéritos online. O inquérito foi ainda enviado à orientação do estudo antes de ser enviado ao painel, para que fossem detetados e corrigidos erros. A segunda fase do estudo teve início no dia 02/04/2015, e terminou no dia 09/04/2015. Os questionários foram enviados aos dezoito membros do painel e a taxa de resposta foi de 100%. A seguinte tabela evidencia os 10 fatores com classificações mais altas (no Anexo I é possível consultar todas as classificações):

Fator Crítico de Sucesso	Média ponderada
Análise de dados automatizada ou semi-automatizada.	8,78
Desenvolver uma abordagem de <i>big data</i> alinhada com os objetivos estratégicos da organização.	8,72
Possuir novos modelos e ferramentas capazes de processar grandes volumes de dados.	8,50
Possuir práticas de <i>Business Intelligence</i> (tais como análise preditiva e <i>big data analytics</i>).	8,39
Tomada de decisão suportada em algoritmos automatizados.	8,39
Capacidade de inovação e adaptação do modelo de negócio.	8,33
Possuir infraestrutura de TI adequada.	8,28
Organização deve permitir e incentivar a experimentação (e.g. testes rápidos e em pequena escala).	8,28
Organização deve ser capaz de implementar soluções <i>big data</i> escaláveis.	8,17
Cultura organizacional orientada para a gestão dos dados.	8,11

Tabela 5 – Lista dos dez fatores com classificações mais altas.

3.4. Terceira Fase

Para as rondas optou-se por realizar uma a três no máximo pois de acordo com Hasson et al (2000) este é o número ideal. Por sua vez a taxa de resposta pretendida a cada ronda foi no mínimo de 70%, uma vez que esta é a taxa mínima apontada pelos mesmos autores para que o estudo se torne relevante.

Relativamente ao nível de consenso, para este estudo adotou-se o valor proposto por Hasson et al (2000) de Kendall W igual a 0,5 representando concordância alta. No entanto caso não seja atingido este valor mas as iterações resultem em valores Kendall W semelhantes, Schmidt (1997) defende que o processo deve ser finalizado. Por outro lado, o valor do coeficiente de correlação de Spearman entre rondas pretende-se que seja o mais próximo possível de 1.

Nesta fase utilizou-se um recurso específico dos inquéritos desenhados na ferramenta online *SurveyMonkey*, que permite criar rankings. Foi então solicitado ao painel que ordenasse os dez fatores, colocando em primeiro lugar o mais importante e em décimo lugar o menos importante.

3.4.1. Terceira Fase – Primeira Ronda

A primeira ronda teve início no dia 14/04/2015, e encerrou no dia 24/04/2015. A taxa de resposta foi de 83%. Nesta fase recorreu-se também à ferramenta SPSS para tratamento dos dados. Para a introdução dos dados nesta ferramenta, os fatores foram ordenados alfabeticamente como evidencia o Anexo II. O valor de Kendall W foi de 0,214 o que indica concordância fraca. A tabela seguinte apresenta os resultados:

Ranking	Fator Crítico de Sucesso	Pontuação
1º	Desenvolver uma abordagem de <i>big data</i> alinhada com os objetivos estratégicos da organização.	8,00
2º	Possuir infraestrutura de TI adequada.	6,60
3º	Possuir novos modelos e ferramentas capazes de processar grandes volumes de dados.	6,53
4º	Cultura organizacional orientada para a gestão dos dados.	5,87
5º	Análise de dados automatizada ou semi-automatizada.	5,73
6º	Organização deve permitir e incentivar a experimentação (e.g. testes rápidos e em pequena escala).	5,47
7º	Possuir práticas de <i>Business Intelligence</i> (tais como análise preditiva e <i>big data analytics</i>).	5,13
8º	Capacidade de inovação e adaptação do modelo de negócio.	4,53
9º	Tomada de decisão suportada em algoritmos automatizados.	4,00
10º	Organização deve ser capaz de implementar soluções <i>big data</i> escaláveis.	3,13

Tabela 6 – Ordenação dos fatores após a 1ª ronda.

No Anexo III é possível verificar os dados estatísticos desta fase.

Nesta fase é ainda dada a hipótese do painel comentar ou justificar o ranking criado. No entanto foi recebido apenas o seguinte comentário:

A opção foi dar relevância aos aspetos do big data alinhados com a estratégia. Sem uma estratégia que reconheça o contributo do big data pouco sentido faria adotar medidas avulso. Depois a seguir aparecem os fatores táticos e os dependentes da infraestrutura, por exemplo não faz sentido pensar em soluções escaláveis sem uma infraestrutura adequada.

3.4.2. Terceira Fase – Segunda Ronda

Tendo em conta o baixo valor do W de Kendall decidiu-se tentar melhorar o mesmo. A segunda ronda teve início no dia 29/04/2015, e encerrou no dia 18/05/2015. Foi partilhado ao painel os resultados da ronda anterior e indicado que o objetivo ao repetir a ronda seria aumentar o valor W de Kendall. A taxa de resposta desceu para 72%. O valor de Kendall W obtido foi de 0,313 o que indica concordância moderada. O valor de Spearman rho foi de 0,952 indicando que apesar do moderado grau de concordância no painel, existiu uma alta convergência de opiniões entre as rondas (Anexo V).

Após esta ronda decidiu-se encerrar a recolha de dados. Esta decisão teve em conta três fatores presentes na literatura que justificam a paragem das rondas: (1) Hasson et al (2000) indicam que as rondas devem parar caso o número de respostas diminua, principalmente quando se trata apenas de um painel; (2) Se a alteração no valor de Kendall W não for significativo, como foi o caso, o estudo deve ser encerrado, uma vez que será retirada pouca informação adicional de rondas futuras (Schmidt, 1997; Habibi et al, 2014); (3) Por fim, Kalaian e Kasim (2012) afirmam que se o valor de correlação Spearman estiver próximo do valor 1, é um indicador que existe já uma forte convergência nas opiniões entre as rondas, ou seja o painel manteve a sua opinião. Desta forma será difícil aumentar o grau de concordância Kendall W. A ordenação final do estudo é a seguinte:

Ranking	Fator Crítico de Sucesso	Pontuação
1º	Desenvolver uma abordagem de <i>big data</i> alinhada com os objetivos estratégicos da organização.	7,77
2º	Possuir infraestrutura de TI adequada.	7,46
3º	Possuir novos modelos e ferramentas capazes de processar grandes volumes de dados.	7,31
4º	Cultura organizacional orientada para a gestão dos dados.	6,31
5º	Análise de dados automatizada ou semi-automatizada.	5,77
6º	Possuir práticas de <i>Business Intelligence</i> (tais como análise preditiva e <i>big data analytics</i>).	4,92
7º	Capacidade de inovação e adaptação do modelo de negócio	4,62
8º	Organização deve permitir e incentivar a experimentação (e.g. testes rápidos e em pequena escala).	4,46
9º	Organização deve ser capaz de implementar soluções <i>big data</i> escaláveis.	3,23
10º	Tomada de decisão suportada em algoritmos automatizados	3,15

Tabela 7 – Ordenação final dos fatores.

No Anexo IV é possível verificar as informações de estatística descritiva da última ronda.

4. Análise e Discussão dos Resultados

Após a recolha de dados foi possível verificar que apesar do painel ser formado por profissionais com conhecimento do tema em estudo registou-se uma concordância moderada nas opiniões relativas ao ranking. Tendo em conta estes valores de concordância, o estudo pode tornar-se um reflexo do *big data*, que ainda parece não reunir muito consenso. Como Steinberg (2014) afirma, apesar do crescente ênfase na importância do *big data*, ainda existe falta de consenso sobre aquilo que é necessário realmente fazer para tirar proveito do mesmo.

Apesar do moderado grau de concordância, é de realçar a forte convergência das opiniões entre as duas rondas. De facto, foi possível observar que não houve alterações nas posições dos cinco primeiros fatores.

Ficou assim patente, a importância do alinhamento das abordagens *big data* nas organizações com os seus objetivos estratégicos. Apesar de na segunda fase do estudo este FCS ter obtido a segunda classificação mais alta, quando foi solicitado ao painel para estabelecer rankings, em ambas as rondas ficou em 1º lugar. De facto para Rijmenam (2013a) a abordagem ao *big data* deve ser encarada como estratégica e não como mais uma disciplina no âmbito de TI. Desta forma é possível definir: objetivos; um eventual ROI; os impactos; a possibilidade de melhorar processos; e outros critérios importantes para obter sucesso num projeto de *big data*.

A segunda posição em ambas as rondas foi ocupada pelo FCS “Possuir infraestrutura de TI adequada”. No inquérito elaborada pela Cisco (2012) as

questões sobre a relação entre *big data* e infraestrutura registaram os seguintes resultados: apenas dois em cada cinco respondentes acreditavam que a sua rede atual estava preparada para *big data*, e um em cinco necessitavam de mais largura de banda. Buhl et al (2013) afirmam que para beneficiar do *big data* é inevitável melhorar a infraestrutura tecnológica da organização, podendo refletir-se em alterações nas aplicações e por conseguinte nos processos de negócio. A infraestrutura TI deve apresentar uma baixa e previsível latência na captura e consultas simples aos dados; ser capaz de lidar com grandes quantidades de volumes transacionados, muitas vezes em ambientes distribuídos; e suportar estruturas de dados flexíveis e dinâmicas (Heisterberg e Verma, 2014).

“Possuir novos modelos e ferramentas capazes de processar grandes volumes de dados” foi o terceiro colocado nas duas rondas. Heisterberg e Verma (2014) afirmam que as organizações devem aumentar a sua preocupação com a capacidade de analisar novos volumes de dados. É neste contexto que surge a necessidade de desenvolver novos modelos para processar a vasta quantidade de dados gerados atualmente.

A quarta posição foi ocupada pelo FCS “Cultura organizacional orientada para a gestão dos dados”. Barton et al (2012) juntamente com uma equipa do MIT Center for Digital Business realizaram um estudo sobre o impacto do *big data*, focando-se no desempenho das organizações orientadas para gestão dos dados. Para tal, realizaram entrevistas estruturadas a executivos de 330 organizações norte americanas. Após a análise das entrevistas, relatórios anuais e informações de fontes independentes foi possível concluir que as

organizações orientadas para gestão dos dados eram 5% mais produtivas e registaram lucros superiores aos concorrentes em 6%. Estes resultados ficaram a dever-se essencialmente à “*utilização do big data para melhorar as previsões, e melhores previsões levam a uma tomada de decisão melhor*” (Barton et al, 2012, p.2).

No último lugar do top 5 ficou a “Análise de dados automatizada ou semi-automatizada”. Bryant et al (2008) indicam que esta análise são técnicas para detetar padrões, identificar anomalias e extrair conhecimento. As novas formas de computação que combinam análise estatística, otimização de processos e inteligência artificial, são capazes de construir modelos estatísticos a partir de grandes conjuntos de dados e inferir como o sistema deve responder a novos dados. Por exemplo a Netflix usa aprendizagem automática no seu sistema de recomendações, ao prever os interesses do cliente comparando o seu histórico de filmes com um modelo estatístico gerado a partir dos hábitos de milhões de outros clientes. Como Bryant et al (2008) indicam “*a análise automatizada ou semi-automatizada de grandes volumes de dados está no centro do big data computing para todos os domínios aplicacionais*” (p.4).

O sexto lugar da ronda final foi ocupado por “Possuir práticas de *Business Intelligence* (tais como análise preditiva e *big data analytics*)”. Desta forma, este FCS subiu uma posição. A Accenture (2011) afirma que as organizações mais perspicazes já reconhecem a importância do *analytics*. Isto deve-se ao *analytics* levar a uma melhor compreensão do cliente e do mercado. Num estudo levado a cabo pela IBM e MIT Sloan Management

Review, as organizações com melhor desempenho possuíam *analytics* eficiente como fator de diferenciação. Desta forma a principal barreira encontrada para ganhar vantagem competitiva do *big data* foi a falta de compreensão em como utilizar *analytics* para melhorar a organização - essencialmente capturar, representar e entregar (Lycett, 2013). O grande tamanho dos conjuntos de dados não significa automaticamente que será possível retirar bastantes informações dos mesmos, apenas que o potencial de aprendizagem é maior. É essencial existirem capacidades e práticas analíticas na organização para tirar as conclusões necessárias (Pease, 2015).

O FCS “Capacidade de inovação e adaptação do modelo de negócio” também subiu uma posição, ficando assim no sétimo lugar do ranking. Para Barton et al (2012) as organizações estão a reunir grandes conjuntos de dados, no entanto em muitos casos não sabem o que fazer com os mesmos. Sendo assim, para transformar estes dados em informação e conhecimento capaz de oferecer vantagem competitiva à organização, será necessário desenvolver novas capacidades e práticas de gestão, e reformular o modelo de negócio da organização. De acordo com a Search CIO (2014) este novo modelo de negócio, independentemente do setor de atuação da organização, deve entender que a capacidade de gestão e mineração de dados será um fator crítico para o sucesso organizacional.

Por outro lado o FCS “Organização deve permitir e incentivar a experimentação (e.g. testes rápidos e em pequena escala)” desceu uma posição entre as duas rondas, ocupando assim o oitavo lugar. Bughin et al (2010) afirmam que no contexto do *big data*, testes rápidos e em pequena

escala são essenciais, uma vez que permitem expandir de forma melhorada soluções para toda a organização. Para Brown et al (2011a) a experimentação no *big data* também leva à identificação de necessidades, expondo variabilidades e possibilitando uma melhoria do desempenho organizacional.

A nona posição do ranking foi ocupada pelo FCS “Organização deve ser capaz de implementar soluções *big data* escaláveis” registando assim uma subida na classificação. Madden (2012) indica que as capacidades de escalabilidade nas soluções focadas no armazenamento e gestão dos dados são essenciais atualmente. Sejam eles dados estruturados ou não estruturados. De facto Verheij (2013) indica que a adoção de soluções *big data* está a ser essencialmente influenciada pelas características de escalabilidade das mesmas. Sendo as bases de dados NoSQL (Not Only SQL) um exemplo de solução capaz de lidar com o *big data*, graças à capacidade de ultrapassar as dificuldades de escalabilidade presentes nas bases de dados relacionais.

Por fim o ranking é encerrado pelo FCS “Tomada de decisão suportada em algoritmos automatizados”, que por sua vez desceu uma posição na classificação. Para Brinkkemper et al (2013) a “explosão” de dados em termos de volume e formato, originou a necessidade de novas abordagens para processar e analisar grandes quantidades de dados em tempo real. Essencialmente através de novos algoritmos ou adaptações, capazes de tornar dados em *insights* utilizáveis e úteis para a organização na sua tomada de decisão.

5. Conclusões

O presente estudo teve como objetivo identificar e ordenar de acordo com o seu grau de importância, os principais fatores críticos de sucesso para que a organização seja capaz de retirar valor do *big data*. A grande motivação na escolha deste tema reside no facto dos dados estarem atualmente em toda a parte – em todos os setores, em todas as economias, em todas as organizações nos seus processos de negócio. Por outro lado esta é uma temática ainda pouco estudada, como Brown et al (2011a) afirmam, o *big data* ainda se encontra numa fase inicial de pesquisa.

Apesar do moderado grau de concordância na última ronda, o estudo empírico evidencia que a ordenação dos cinco fatores considerados como mais importantes, apresentou bastante coerência através do coeficiente de correlação de Spearman, uma vez que ocuparam exatamente os mesmos lugares em ambas as rondas. No entanto o moderado grau de concordância denota ainda alguma discordância sobre o tema em estudo. Como afirma Langer (2014), *big data* reúne ainda pouco consenso entre profissionais. Esta falta de consenso está ainda evidente nos resultados do inquérito levado a cabo pela Treasure Data (2012) sobre o investimento e adoção de soluções *big data*.

Desta forma é possível concluir que o estudo reflete a visão de Bughin et al (2010) quando estes afirmam que às organizações e gestores não deve chegar apenas compreender as tendências tecnológicas. É necessário também uma preocupação estratégica para adaptar processos e as estruturas

organizacionais de forma a responder às novas exigências e ser capaz de retirar valor das transformações. As organizações e decisores precisam abordar desafios consideráveis se pretenderem capturar todo o potencial do *big data*.

A grande limitação do presente estudo foi o grau de concordância obtido. O resultado do valor de Kendall W foi de 0,313 indicando consenso moderado. Este valor pode refletir a diversidade de opiniões presentes no painel, sobre uma temática ainda pouco estudada. Enquanto que uma ronda adicional na 3ª fase poderia aumentar este valor, foi ainda tido em conta a crescente dificuldade em obter respostas e a fadiga do painel relativamente ao estudo. Este facto refletiu-se pelo aumento de lembretes enviados na 2ª ronda a solicitar as respostas do painel.

No entanto o estudo pode fornecer uma base para os profissionais assim como para pesquisas futuras. Seria interessante perceber e comparar as diferenças entre os fatores críticos de sucesso para retirar valor do *big data* e para retirar valor de dados estruturados.

A partir dos resultados obtidos poderia ainda ser benéfico tentar desenvolver *frameworks*, para garantir que os FCS identificados no estudo são concretizados e quais as métricas para medir o sucesso.

Por fim seria importante perceber como está a ser abordado o *big data* em Portugal e perceber se os fatores presentes no estudo divergem muito daqueles que a realidade das organizações portuguesas identifica como sendo os mais importantes para retirar valor do *big data*.

6. Referências Bibliográficas

- Accenture. (2011). Accenture Technology Vision 2011.
- Agrawal, D., Bernstein, P., Bertino, E., Davidson, S., Dayal, U., Franklin, M., ... Widom, J. (2012). Challenges and opportunities with big data.
- Ahmed, G. (2014). Critical success factors and why CSFs are important for business success.
- Baesens, B., Bapna, R., Marsden, J. R., Vanthienen, J., & Zhao, J. L. (2014) Transformational Issues of Big Data and Analytics in Networked Business. *MIS Quarterly*.
- Barton, D., Brynjolfsson, E., Davenport, T. H., McAfee, A., & Patil, D. J. (2012). Big Data. *The management revolution. Harvard Bus Rev*, 90(10), 59-68.
- Bayer, L. R., & Wittink, D. R. (1994). The measurement imperative. *Marketing Research*, 6(4), 14-23.
- Belassi, W., & Tukel, O. I. (1996). A new framework for determining critical success/failure factors in projects. *International journal of project management*, 14(3), 141-151.
- Boone, H. N., & Boone, D. A. (2012). Analyzing likert data. *Journal of Extension*, 50(2), 1-5.
- Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*, 15(5), 662-679.

- Brinkkemper, S., Fotaki, G., Meijer, D., & Spruit, M. (2013). Exploring Big Data Opportunities for Online Customer Segmentation. *Mycotoxin Research*, 6, 100.
- Brown, B., Bughin, J., Byers, A.H., Chui, M., Dobbs, R., Manyika, J. & Roxburgh, C. (2011a). Big data: The next frontier for innovation, competition, and productivity.
- Brown, B., Chui, M., & Manyika, J. (2011b). Are you ready for the era of 'big data'. *McKinsey Quarterly*, 4, 24-35.
- Bryant, R., Katz, R. H., & Lazowska, E. D. (2008). Big-data computing: creating revolutionary breakthroughs in commerce, science and society.
- Bughin, J., Chui, M., & Manyika, J. (2010). Clouds, big data, and smart assets: Ten tech-enabled business trends to watch. *McKinsey Quarterly*, 75–86.
- Buhl, H., Heidemann, J., Moser, F., & Röglinger, M. (2013) Big Data - A Fashionable Topic with(out) Sustainable Relevance for Research and Practice?, *Business & Information Systems Engineering*, 5(2), 65-69.
- Bullen, C. V., & Rockart, J. (1981). A Primer on Critical Success Factors. *Center for Information Systems Research Working Paper*, (69).
- Caldeira, M. M., & Ward, J. M. (2003). Using resource-based theory to interpret the successful adoption and use of information systems and technology in manufacturing small and medium-sized enterprises. *European journal of information systems*, 12(2), 127-141.
- Cisco. (2012). Big data and beyond: Data in motion.

- DeRoos, D., Deutsch , T., Eaton, C., Lapis, G., & Zikopoulos, P. C. (2012). *Understanding big data: Analytics for enterprise class hadoop and streaming data*. New York et al: McGraw-Hill.
- Deng, T., Keil, M., & Lee, H. K. (2013). Understanding the most critical skills for managing IT projects: A Delphi study of IT project managers. *Information & Management*, 50(7), 398-414.
- Fattah, A. (2014). Going Beyond Data Science Toward an Analytics Ecosystem : Part 1. *IBM Data Magazine*, 3–5.
- Fawcett, S. E., & Waller, M. A. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *Journal of Business Logistics*, 34(2), 77–84.
- Ferguson, M. (2013). Enterprise Information Protection - The Impact of Big Data.
- Gartner. (2012a). Gartner Says Big Data Creates Big Jobs: 4.4 Million IT Jobs Globally to Support Big Data by 2015, [Press Release]. Retirado de <http://www.gartner.com/newsroom/id/2207915>
- Gartner. (2012b). Gartner Says Big Data Disruptions Can Be Tamed With Enterprise Architecture, [Press Release].
- Habibi, A., Sarafrazi, A., & Izadyar, S. (2014). Delphi Technique Theoretical Framework in Qualitative Research. *The International Journal Of Engineering And Science (IJES)*, 3(4), 08-13.
- Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of advanced nursing*, 32(4), 1008-1015.

- Heisterberg, R., & Verma, A. (2014). *Creating Business Agility*. Wiley.
- IBM. (2013a). What is Big Data. Disponível em <http://www.ibm.com/big-data/us/en/>
- IBM. (2013b). The four V's of big data.
- IBM. (2014). "What is a Data Scientist". Disponível em <http://www-01.ibm.com/software/data/infosphere/data-scientist/>
- Jacoby, J., & Matell, M. S. (1971). Is There an Optimal Number of Alternatives for Likert Scale Items? Study. *Educational and psychological measurement*, 31, 657-674.
- Kalaian, S. A., & Kasim, R. M. (2012). Terminating sequential Delphi survey data collection. *Practical Assessment, Research & Evaluation*, 17(5), 2.
- Kandavel, D., & Sundaresan, B. (2014). Big Languages for Big Data.
- Lajara, T., Maçada, A., & Brinkhues, R. (2013). Investimento em Ti e sua influência no desempenho organizacional: Através da governança da informação e capacidades de gestão da informação, moderado pelo big data.
- Langer, G. (2014). Growing doubts about big data.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*.
- Lohr, S. (2012). The age of big data. *New York Times*, 11.
- Lycett, M. (2013). "Datafication": making sense of (Big) data in a complex world. *European Journal of Information Systems*, 22, 381–386.

- Lynch, R. (2003). *Corporate Strategy*. Harlow. *Financial Times/Prentice Hall*.
- Meyer, E., Schroeder, L., & Taylor, L. (2014). Emerging practices and perspectives on Big Data analysis in economics: Bigger and better or more of the same?. *Big Data & Society*, 1(2).
- Munshi, J. (2014). A method for constructing Likert scales.
- Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. *Information & Management*, 42(1), 15-29.
- Pal, R., & Torstensson, H. (2011). Aligning critical success factors to organizational design: A study of Swedish textile and clothing firms. *Business Process Management Journal*, 17(3), 403-436.
- Pease, A. (2015). Estatística na era do *big data* e dos *data scientists*.
- Rijmenam, M. (2013a). The big data roadmap to a winning big data strategy.
- Rijmenam, M. (2013b). Finding the right sponsor for your big data project is vital.
- Rockart, J. F. (1982). *The changing role of the information systems executive: a critical success factors perspective*. Boston: Massachusetts Institute of Technology.
- Sas. (2011). Survey: Strategic data management yields competitive advantage. [Press release].
- Schmidt, R. C. (1997). Managing Delphi surveys using nonparametric statistical techniques. *Decision Sciences*, 28(3), 763-774.

- Search CIO. (2014). From data gathering to competitive strategy: the evolution of big data.
- Steinberg, D. (2014). 3 Ways companies can better use big data in 2014.
- Treasure Data. (2012). Big Data Survey Results.
- Turner, M. K. (2014). Three Keys to IT System Success: People, Process, Technology.
- Varonis Systems. (2012). Big data and infosecurity: Research Report.
- Verheij, B. A. (2013). *The process of big data solution adoption* (Doctoral dissertation, TU Delft, Delft University of Technology).
- Walker, D. H. T. (1995). *An investigation into factors that determine building construction time performance*.

7. Anexos

Anexo I

Fator Crítico de Sucesso	Média ponderada
Análise de dados automatizada ou semi-automatizada.	8,78
Desenvolver uma abordagem de <i>big data</i> alinhada com os objetivos estratégicos da organização.	8,72
Possuir novos modelos e ferramentas capazes de processar grandes volumes de dados.	8,50
Possuir práticas de <i>Business Intelligence</i> (tais como análise preditiva e <i>big data analytics</i>).	8,39
Tomada de decisão suportada em algoritmos automatizados.	8,39
Capacidade de inovação e adaptação do modelo de negócio.	8,33
Possuir infraestrutura de TI adequada.	8,28
Organização deve permitir e incentivar a experimentação (e.g. testes rápidos e em pequena escala).	8,28
Organização deve ser capaz de implementar soluções <i>big data</i> escaláveis.	8,17
Cultura organizacional orientada para a gestão dos dados.	8,11
Facilitar acesso dos <i>stakeholders</i> aos dados.	8,00
Respeitar segurança e privacidade dos dados (clientes preocupados com a segurança dos seus dados).	8,00
<i>Enterprise Architecture</i> estar preparada para <i>big data</i> .	7,94
Existir recursos humanos com capacidade para abordar <i>big data</i> .	7,83
Reduzir custos de gestão de dados mas aumento da velocidade de tratamento.	7,44
Existir alta qualidade dos dados.	7,44
Realizar investimentos e políticas orientadas para o futuro.	7,39
Evitar acumulação de dados em silos departamentais.	7,33
Obter <i>sponsorship</i> para projetos <i>big data</i> .	6,94
Organização deve ter plataforma de experiência do cliente multi-canal.	6,72

Tabela 8 – Lista completa das classificações dos fatores.

Anexo II

FCS	Descrição
1	Análise de dados automatizada ou semi-automatizada.
2	Capacidade de inovação e adaptação do modelo de negócio.
3	Cultura organizacional orientada para a gestão dos dados.
4	Desenvolver uma abordagem do <i>big data</i> alinhada com os objetivos estratégicos da organização.
5	Organização deve permitir e incentivar a experimentação (e.g. testes rápidos e em pequena escala).
6	Organização deve ser capaz de implementar soluções big data escaláveis.
7	Possuir infraestrutura de TI adequada.
8	Possuir novos modelos e ferramentas capazes de processar grandes volumes de dados.
9	Possuir práticas de <i>Business Intelligence</i> (tais como análise preditiva e <i>big data analytics</i>).
10	Tomada de decisão suportada em algoritmos automatizados.

Tabela 9 – Ordenação alfabética dos fatores.

Anexo III

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
FCS1	15	5.2667	2.01660	2.00	9.00
FCS2	15	6.4667	2.64215	1.00	9.00
FCS3	15	5.1333	3.41983	1.00	10.00
FCS4	15	3.0000	2.80306	1.00	10.00
FCS5	15	5.5333	2.41622	2.00	9.00
FCS6	15	7.8667	2.19957	3.00	10.00
FCS7	15	4.4000	2.97129	1.00	10.00
FCS8	15	4.4667	2.66905	1.00	10.00
FCS9	15	5.8667	2.44560	2.00	10.00
FCS10	15	7.0000	2.50713	1.00	10.00

Tabela 10 – Estatística descritiva da 3ª fase, 1ª ronda.

Test Statistics

N	15
Kendall's W	.214
Chi-Square	28.862
df	9
Asymp. Sig.	.001

Tabela 11 – Resultado do coeficiente W de Kendall 1ª ronda.

Anexo IV

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
FCS1	13	5.2308	2.48843	1.00	10.00
FCS2	13	6.3846	2.87340	1.00	10.00
FCS3	13	4.6923	3.06552	1.00	9.00
FCS4	13	3.2308	2.61896	1.00	9.00
FCS5	13	6.5385	1.80810	4.00	10.00
FCS6	13	7.7692	2.42053	3.00	10.00
FCS7	13	3.5385	2.69615	1.00	10.00
FCS8	13	3.6923	1.49358	1.00	7.00
FCS9	13	6.0769	2.98501	1.00	10.00
FCS10	13	7.8462	1.77229	4.00	10.00

Tabela 12 – Estatística descritiva da 3ª fase, 2ª ronda.

Test Statistics

N	13
Kendall's W	.313
Chi-Square	36.642
df	9
Asymp. Sig.	.000

Tabela 13 – Resultado do coeficiente W de Kendall 2ª ronda.

Anexo V

Correlations

		Ronda_1	Ronda_2
Spearman's rho	Correlation Coefficient	1.000	.952
	Ronda_1 Sig. (2-tailed)	.	.000
	N	10	10
	Correlation Coefficient	.952	1.000
	Ronda_2 Sig. (2-tailed)	.000	.
	N	10	10

Tabela 14 – Resultado do coeficiente de correlação Spearman rho.