



NADINE
CÔRTE-REAL

BIG DATA & ANALYTICS

O poder de transformar dados em inteligência
e o impacto na competitividade empresarial

INFLUÊNCIA

Introdução	6
1. A revolução de big data e o impacto na tomada de decisão das organizações	18
2. Governança de dados: o caminho necessário para garantir a saúde dos dados	46
3. Gestão de inovação e transformação digital alavancada por big data analytics e inteligência artificial	82
4. A estratégia de big data analytics na inovação e transformação digital dos processos da organização	100
5. Impacto da inteligência artificial na sociedade e nas empresas	114
6. Adoção e implementação de soluções de IA de forma responsável	156
7. Conclusões e considerações finais	206
Glossário	212
Agradecimentos	218
Recomendações de leitura	220

INTRODUÇÃO

Sempre tive vontade de escrever um livro sobre algo útil para a sociedade. Os livros assumem-se como uma ótima forma de difundir conhecimento e potenciar as mentes de quem os lê. Como diria o empreendedor Naval Ravikant: «Para escrever um grande livro, precisas primeiro de te tornar o livro.»

Considero-me uma apaixonada pela área de gestão de informação e inovação, pois acredito que a correta utilização deste conhecimento permite melhorar a vida das empresas e das pessoas através da otimização do seu tempo.

No meu percurso profissional e académico fui-me deparando com vários problemas relacionados com a gestão de informação nas empresas. Ao longo de 14 anos de investigação e participação em projetos de gestão de informação, adquiri conhecimentos sobre as melhores práticas quer na área de governação, quer na área de análise de informação. Recentemente terminei o doutoramento em Big Data Analytics e percebi que, mesmo tendo contribuído para aumentar o conhecimento nesta área, existe muita informação que deveria estar consolidada para se poder verdadeiramente tirar partido deste tipo de iniciativas nas empresas. Através dos cursos que leciono, conheci muitas pessoas intrigadas e perdidas na forma de utilizar este ativo estratégico que são os dados.

Este livro representa a consolidação de conhecimentos que gostaria de ter tido acesso ao longo dos anos de experiência a trabalhar nesta área e algumas das necessidades trazidas pelos meus alunos nas aulas. Pretendo, assim, que o leitor possa ganhar o tempo que perdi a solucionar os problemas e o transforme em investimento para si a nível pessoal ou profissional.

CONTEXTO

Atualmente, vivemos na era digital da informação onde o saber olhar para a informação certa no momento certo é crucial na tomada de decisão dos indivíduos e das empresas. Nunca existiu, em momento algum da história, tantos dados a serem produzidos, trocados e partilhados a um ritmo exponencial e alucinante. A este fenómeno chamamos «big data». Nos tempos que correm, big data constitui-se como um verdadeiro tsunami de dados, uma área em grande desenvolvimento que, quando governada e suportada por tecnologia inovadora — nomeadamente big data analytics e inteligência artificial (IA) —, é capaz de oferecer formas inteligentes de retirar valor dos mesmos.

Nunca, em momento algum, foi tão difícil selecionar informação fidedigna, pois são tantas as fontes de desinformação como as de informação. Todos nós produzimos dados através de pequenos atos do dia a dia, seja um breve acesso a uma rede social ou uma mera consulta no *Google*. Os nossos comportamentos no mundo digital alimentam os famosos algoritmos criados por empresas como a Meta (proprietária da Facebook) ou a Google, sendo estes usados para captar padrões e incentivar o consumo personalizado de informação. Nas empresas, todos os dias se geram e são transacionados

dados para suportar processos de negócio. Toda esta informação, se devidamente consolidada e preparada, é possível de ser analisada. Deste modo, a informação criada em torno da sociedade digital tem-se tornado cada vez mais um ativo estratégico para as empresas. É fundamental governar os dados de forma estratégica e torná-los inteligentes para obtenção de vantagens competitivas. Contudo, apesar da adoção e implementação destas tecnologias de big data analytics e inteligência artificial ser cada vez mais imperativa não é linear qual o melhor caminho a seguir. As empresas desconhecem as melhores formas de introduzir estas tecnologias para inovar os seus processos de negócio.

Mais do que uma sociedade de informação são precisas pessoas informadas de como potenciá-la e geri-la de forma efetiva. Para as iniciativas de gestão de informação serem bem-sucedidas, os *stakeholders*¹ de negócio já não podem delegar todas as iniciativas para o IT² e necessitam de participar ativamente. Os responsáveis de negócio precisam de entender onde e como alavancar os dados, explorando fontes e utilizando técnicas para otimizar os seus processos de negócio. Deste modo, é possível promover oportunidades de monetização dos dados e criar vantagens competitivas.

Aqui encontrará conselhos práticos. O conteúdo foca-se em conceitos, princípios e metodologias que podem ser aplicadas a qualquer indústria ou contexto tecnológico. É orientado para o negócio e tem como objetivo apresentar um conjunto de boas práticas e recomendações a gestores de negócio responsáveis, ou elementos de equipa em iniciativas de gestão de informação, quer seja na perspetiva analítica quer seja na de governação.

^{1,2} Ver termos no glossário.

É importante notar que este é um guia prático, mas cada empresa deve realizar uma análise crítica da transposição destas mesmas práticas e orientações para a sua realidade. Não existem abordagens únicas, pois os contextos empresariais também não são únicos. A principal ideia que é transmitida ao longo destas páginas é que a realidade de cada empresa é a base para a implementação de iniciativas de gestão de informação. A tecnologia só é uma necessidade para as empresas porque serve de apoio ao cumprimento de objetivos de negócio. Assim, as iniciativas e as estratégias de dados devem nascer das necessidades e prioridades de negócio existentes e fluir sempre em torno dos objetivos estratégicos da empresa.

Este livro encontra-se organizado em seis capítulos essenciais para a perceção do ecossistema de dados e big data de forma geral, bem como para a compreensão de programas de governação de dados e transformação digital (com particular atenção em tecnologias de big data analytics e inteligência artificial). A combinação destas peças permite abordar de forma transversal as problemáticas da gestão de informação num contexto empresarial.

Não existem pré-requisitos para a sua leitura, não necessitando também de ser lido de forma corrida, podendo ser consultado de acordo com a necessidade do leitor.

CAPÍTULO 1 — A REVOLUÇÃO DE BIG DATA E O IMPACTO NA TOMADA DE DECISÃO DAS ORGANIZAÇÕES

Neste capítulo inicial são introduzidos os principais conceitos necessários para compreender o fenómeno do big data e o seu ecossistema. Para demonstrar de forma prática o impacto do big data na sociedade, e de modo mais particular nas empresas, são

apresentados vários casos de sucesso de organizações presentes em vários setores.

De forma a garantir o sucesso das iniciativas de big data, torna-se necessário existir uma mudança de mentalidade nas organizações. Algumas destas medidas são apresentadas neste capítulo sendo que constituem a base para a extração de valor dos dados. Ao percebermos as características e a importância dos dados, podemos facilmente perceber a necessidade de investimento em temáticas abordadas nos capítulos seguintes.

CAPÍTULO 2 — GOVERNAÇÃO DE DADOS: O CAMINHO NECESSÁRIO PARA GARANTIR A SAÚDE E DOS DADOS

Muitas vezes, quando se fala em dados e/ou big data e a necessidade de retirar valor dos mesmos, existe uma falsa percepção de que basta investir em pessoas que saibam programar e criar algoritmos sofisticados para resolver problemas de análise de dados.

Para gerir e retirar valor de dados e em particular de big data, mais do que ter iniciativa isolada ou departamental, deve existir um programa corporativo de governação de dados que permita oferecer uma visão única dos dados à organização e monetizar oportunidades de negócio com os mesmos. Com este tipo de visão, a organização consegue ter confiança para usar os dados na tomada de decisão de forma a atingir os seus objetivos estratégicos.

Este capítulo demonstra a necessidade de definição e implementação de uma estratégia de governação dos dados. Abordam-se as principais preocupações que as empresas devem considerar (como a qualidade e segurança dos dados), e dão-se a conhecer as principais ferramentas e metodologias que podem ser usadas na sua implementação.

CAPÍTULO 3 — GESTÃO DE INOVAÇÃO E TRANSFORMAÇÃO DIGITAL ALAVAN-CADA POR BIG DATA ANALYTICS E INTELIGÊNCIA ARTIFICIAL

Após compreender o papel basilar da governação dos dados na gestão da informação corporativa, prossegue-se para o domínio da gestão de inovação. Ao saber que dados existem na organização, como podem ser utilizados e qual a sua qualidade, a empresa cria as condições-base para potenciar um ecossistema de inovação e transformação digital. As empresas querem cada vez mais investir em inovação de forma a lidar com a incerteza (por exemplo, na evolução do seu modelo de negócio para novos produtos e serviços, dada a oferta dos seus concorrentes, e no desenvolvimento tecnológico, entre outros fatores externos). Esta incerteza está cada vez mais presente nos dias que correm, dadas as crescentes pressões regulamentares no mercado empresarial e a agilidade dos concorrentes onde as empresas operam.

Concretamente, neste capítulo abordamos a importância da tecnologia avançada de big data analytics e inteligência artificial nos programas de inovação e transformação digital das organizações. As tecnologias de big data analytics combinadas com inteligência artificial são cada vez mais utilizadas na realidade empresarial. Este capítulo visa reforçar a importância desta relação simbiótica. Aqui apresentamos uma metodologia de gestão de inovação bem como um conjunto de boas práticas para garantir a criação de produtos inteligentes, capazes de responder às necessidades mais críticas do negócio e da organização.

CAPÍTULO 4 — A ESTRATÉGIA DE BIG DATA ANALYTICS NA INOVAÇÃO E TRANSFORMAÇÃO DIGITAL DOS PROCESSOS DA ORGANIZAÇÃO

Para quem está a iniciar-se na jornada de data analytics (e em particular big data) ou pretende redefinir/validar a sua estratégia, este

capítulo oferece uma visão global de como a empresa deve estruturar as iniciativas e avaliá-las para melhorar as capacidades analíticas da organização. Além disso, apresenta ferramentas de avaliação de maturidade tecnológica (particularmente a da TDWI³), de gestão deste tipo de projetos, bem como fatores e riscos a considerar neste percurso.

CAPÍTULO 5 — IMPACTO DA INTELIGÊNCIA ARTIFICIAL NA SOCIEDADE E NAS EMPRESAS

A inteligência artificial está em todo o lado. Se olharmos à nossa volta e para os nossos hábitos pessoais, são vários os exemplos que conseguiríamos destacar onde a IA está presente (por exemplo, telemóveis, eletrodomésticos, automóveis, redes sociais, entre muitos outros). Neste capítulo é apresentado em detalhe o impacto da massificação de IA na sociedade e nas empresas, através de exemplos de aplicação em várias indústrias e o seu potencial de utilização, tendo em conta os vários tipos de IA. São desmistificadas algumas das ideias que se têm vindo a criar em torno de IA (por exemplo, que irá manipular o mundo ou criar desemprego generalizado) e são expostos alguns dos principais desafios futuros (como *upskill* e *reskill* de pessoas, gestão de conhecimento, entre outros), que devem ser considerados quando se pensa em investir nesta tecnologia inovadora.

Em particular, é apresentada uma metodologia referenciada e comprovada de gestão deste tipo de projetos (CRISP DM⁴) para garantir o impacto positivo nas empresas e a consequente extração de valor das mesmas.

³ TDWI – The Datawarehouse Institute é uma entidade que oferece a empresas e indivíduos um conjunto de serviços de formação técnica e de negócio e que realiza investigação em várias áreas tecnológicas.

⁴ CRISP DM – Cross Industry Standard Process for Data Mining – É uma metodologia de referência utilizada para desenvolvimento de projetos de Machine Learning.

CAPÍTULO 6 — ADOÇÃO E IMPLEMENTAÇÃO DE SOLUÇÕES DE IA DE FORMA RESPONSÁVEL

Com a crescente importância e adoção de IA por parte das empresas, começam a surgir cada vez mais preocupações de como gerir este tipo de soluções de forma responsável. São várias as escolhas que uma empresa tem de fazer no desenvolvimento de sistemas. Neste processo surgem questões associadas à ética, justiça, transparência e responsabilidade das soluções. Embora exista muita investigação, dada a rapidez da evolução das tecnologias e a maneira como se relacionam com a sociedade, existe um espaço considerável que gera dúvida e debate. A IA deve focar-se na criação de benefícios para as pessoas e a sociedade. Garantir uma IA ética implica não só os responsáveis pela conceção, implementação e monitorização de sistemas de IA, mas também o envolvimento de toda a comunidade. Através da utilização de IA devemos ter a ambição de superar o ser humano no que diz respeito à ética, à transparência e à responsabilidade, contribuindo para um mundo mais igualitário e inclusivo. Assim, o último capítulo deste livro aborda o conceito emergente de «Inteligência Artificial Responsável e Confiável». Numa ótica de implementação, neste capítulo são apresentados os seus princípios, melhores práticas internacionais, orientações e *frameworks* de IA em várias áreas (como estruturas organizacionais que devem existir, considerações a ter no desenvolvimento de soluções, etc.) e algumas ferramentas de monitorização de IA.

Em suma, este livro pretende apresentar as principais capacidades relativas aos dados (armazenar, gerir, transformar, planear e monetizar gerando valor) em que as empresas devem investir de forma a potenciar e retirar o máximo de valor de negócio para os

seus processos. Não se trata de uma obra científica, mas um livro orientador, de apoio a projetos de governação, gestão e exploração de informação com soluções de IA Responsável, devendo ser consultado de forma sistemática e pedagógica.



Figura 1 — Organização e lógica de evolução dos capítulos do livro.

QUEM DEVE LER O LIVRO

Este livro destina-se principalmente a analistas e gestores que participem em iniciativas de gestão integrada de informação e transformação digital nas suas empresas. De modo geral, é neste tipo

de cultura organizacional orientada para os dados que normalmente nasce o interesse pela exploração de dados e em particular de big data. Não obstante, o livro é feito para o público em geral que tenha interesse por esta temática, cada vez mais emergente na sociedade em que vivemos.

GANHOS PARA O LEITOR

Ao ler este livro pretende-se que os seus leitores adquiram o conhecimento essencial sobre a gestão de informação, transformação digital e a extração de valor através de ferramentas de análise de dados, como big data analytics e inteligência artificial. Com base neste conhecimento, é possível adquirir capacidade de:

1. Observar uma realidade empresarial e identificar problemas de negócio que sejam passíveis de ser resolvidos com iniciativas de gestão de informação, inovação, analytics avançada ou inteligência artificial;
2. Gerir, potenciar a combinação de fontes de dados para responder a problemas de negócio e otimizar processos de negócio;
3. Analisar o estado de maturidade da empresa numa perspetiva analítica e de governação;
4. Definir estratégias e *frameworks* de governação de dados, inovação e ferramentas como big data analytics e inteligência artificial de forma responsável e ágil;
5. Compreender os vários tipos de inteligência artificial e abordagens de análise de dados de forma a perceber como podem ser aplicados nas problemáticas empresariais;
6. Alavancar a adoção de ferramentas de big data analytics e inteligência artificial, considerando os principais fatores de

sucesso e ultrapassar eventuais obstáculos (por exemplo, demonstrar o seu valor, obter financiamento, obter suporte da gestão de topo, etc.);

7. Promover uma cultura analítica orientada para os dados e transformação digital.

**A REVOLUÇÃO DE BIG
DATA E O IMPACTO NA
TOMADA DE DECISÃO DAS
ORGANIZAÇÕES**



Na era de big data, a maioria das empresas, independentemente do seu setor, tamanho ou localização, tem de lidar com volumes de dados cada vez maiores. Apesar de ser um desafio gerir tanta informação, o big data constitui-se como um ativo estratégico escondido nas organizações. O conhecimento que deriva do big data gera valor efetivo quando utilizado, ao oferecer suporte crucial à tomada de decisão empresarial e potenciar a criação de vantagens competitivas. Mas como se obtém esta vantagem competitiva?

Apesar das tecnologias que exploram big data (usualmente designadas big data analytics) serem consideradas como o próximo passo para a inovação, a boa maneira de extrair valor das mesmas permanece relativamente desconhecida. Estas ferramentas permitem armazenar e partilhar conhecimento nas áreas de negócio de forma local ou global, o que leva a uma maior eficiência operacional.

Atualmente, as empresas que investem em ferramentas de big data analytics para suportar os seus processos críticos de negócio tendem a ter ganhos de produtividade de 5% e lucros de 6% acima dos seus concorrentes. Por este motivo, muitas empresas ao nível mundial encontram-se a investir fortemente neste tipo de tecnologias.

O caminho para a obtenção de vantagens competitivas com recurso a ferramentas de big data analytics não é de todo linear e deve ser definido para cada empresa de acordo com o seu contexto. De modo geral, esse caminho passa por construir o ambiente tecnológico, com as fontes de informação críticas, capaz de dar suporte aos processos-chave da organização, de forma atempada, para que a empresa possa tomar as decisões certas no tempo certo. Deve, também, existir uma metodologia de implementação de big data alinhada com o plano estratégico da organização. O objetivo desta ação tem por base obter a capacidade de perceber, adquirir, processar, armazenar e analisar dados, convertendo-os em conhecimento.

O big data e as suas tecnologias associadas representam uma verdadeira oportunidade para as empresas, pois são vários os seus benefícios:

1. Redução de custos — Apesar de o investimento inicial ser substancial, as soluções de armazenamento de dados (por exemplo, *Hadoop*, *Cloud*¹) revelam ter um custo inferior;
2. Melhoria de produtos e serviços — Ao extrair conhecimento das tecnologias de big data analytics e utilizá-lo nas suas rotinas permite otimizar os processos de negócio (por exemplo, conhecer as preferências dos clientes e propor produtos personalizados de acordo com o seu perfil de compra);
3. Vantagem competitiva — Por consequência da melhoria de produtos e serviços através da customização, as empresas tendem a obter vantagens competitivas no mercado onde se inserem;
4. Melhoria da tomada de decisão e agilidade organizacional — Ao sustentar as suas decisões com mais informação de várias

¹ Ver termo no glossário.

fontes, podem tomar-se decisões mais céleres, precisas e menos arriscadas.

Desta forma, a jornada do big data revela-se uma verdadeira oportunidade para as empresas. Apesar de não existirem dúvidas em relação aos seus benefícios, tais valias devem ser ponderadas pelos riscos de utilizar big data. Devido às características do big data (volume, variedade, velocidade e veracidade), existem novos perigos que devem ser mitigados (como a partilha inapropriada de dados e a fraca qualidade dos dados recolhidos) através de uma estratégia ao nível organizacional. Tal deve ser feito de forma faseada, estabelecendo prioridades, de modo que seja possível obter benefícios e gerir os riscos associados. Ao longo do livro, iremos abordar estratégias e fatores que podem ser considerados para maximizar a extração de valor para o negócio.

Neste capítulo introdutório, começamos por perceber o fenómeno do big data e as suas dimensões. Numa vertente mais tecnológica, é explicado o seu ecossistema e o tipo de plataformas e ferramentas que podem suportar a existência de big data nas empresas. De seguida, enumeram-se as orientações estratégicas que devem ser seguidas para garantir o êxito das iniciativas ligadas aos dados. Por fim, alguns casos de sucesso de utilização de big data em empresas de referência são apresentados.

1.1. O QUE É E COMO IDENTIFICAR BIG DATA

Antes de nos debruçarmos sobre o conceito de big data, é necessário entender o conceito-base de dados. Os dados podem ser definidos como uma representação de factos não processados ou

não relacionados e constituem-se como a matéria-prima para construirmos peças de informação.

Apesar de a maioria das pessoas pensar que big data é um fenómeno recente, a verdade é que, de acordo com o *Google Trends*, o interesse por big data começou a aumentar a partir de 2011. Ainda assim, o termo «big data» foi utilizado pela primeira vez em 1997, num artigo escrito por investigadores da NASA. Neste artigo, big data foi definido como «grandes conjuntos de dados que colocam em causa a capacidade de memória». Desde então várias definições surgiram. Mas, hoje em dia, o que é considerado big data?

Big data assume-se como um fenómeno vetorial. Tipicamente, existe a ideia de que é apenas um grande volume de dados, mas é muito mais do que isso.

De acordo com a McKinsey, o big data assume-se como a fronteira para a inovação, competitividade e produtividade. São dados que, pela sua escala, distribuição, diversidade e/ou celeridade, implicam recorrer à utilização de arquiteturas técnicas e analíticas que permitam extrapolar análises através de novas fontes (como a utilização de informação proveniente de redes sociais) com valor para o negócio. Para tirar partido de big data, torna-se essencial investir em soluções eficazes a processar grandes volumes de dados e capazes de integrar diversos tipos de dados (estruturados e não estruturados), garantido o seu controlo de qualidade.

Big data engloba a combinação de volume, velocidade, variedade e veracidade de dados (os chamados 4 V). Estas dimensões, se bem combinadas, podem gerar valor tangível e intangível para a organização.

Volume + Velocidade + Variedade + Veracidade = Valor

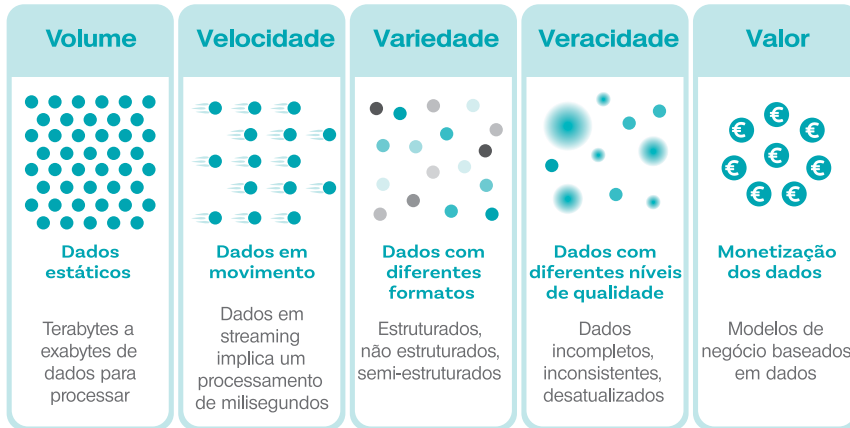


Figura 1.1 — Dimensões de big data (adaptado).

De seguida, são apresentadas em maior detalhe as várias dimensões para que seja claro quais os requisitos que podem ser necessários numa organização.

1.1.1. Big data é velocidade

A informação anda a ser produzida a um ritmo alucinante nunca antes visto. Alguns investigadores acreditam que 90% da informação mundial tenha sido criada desde 2018. Estima-se, igualmente, que a cada dois anos o volume de dados, ao nível mundial, duplique o seu tamanho. Vivemos cada vez mais mergulhados em dados. Mas de que quantidade de dados estamos a falar? Alguns estudos ajudam-nos a dimensionar este fenómeno.

Começemos pelo volume de dados que é gerado por dia no mundo. Num minuto, muita coisa pode acontecer na Internet. Com a crescente utilização de várias plataformas, as estatísticas

de utilização aumentam todos os anos a um ritmo exponencial. Atualmente, o *Facebook* conta com 1.6 milhões de utilizadores a acederem por minuto. Já no *Instagram*, 695 mil fotos são publicadas a cada minuto. No *WhatsApp*, cerca de 69 mil mensagens são enviadas por minuto. Estes são apenas alguns exemplos das várias plataformas que contribuem para esta explosão de dados.

Quando uma organização quer explorar dados das várias plataformas digitais disponíveis, tem de seleccionar ferramentas capazes de integrar essa informação de forma igualmente rápida. Existem atualmente no mercado plataformas que já possuem APIs (*Application Programming Interface*)² para extrair e explorar este tipo de informação (como *ScraperAPI* e *Parsehub*).

Esta é a maneira mais simples de ter acesso a este tipo de dados. A outra opção é construir um programa que o faça. Esta opção em particular apenas é escolhida quando se pretende obter dados de plataformas muito específicas que não possuem APIs.

1.1.2. Big data é volume

Com o volume de dados a aumentar rapidamente, as empresas passam a ter um grande desafio que se pode revelar uma grande oportunidade: ser capazes de olhar e seleccionar a informação que sustenta os seus processos de negócio para obtenção de vantagem competitiva.

Feitas as contas, todas as plataformas utilizadas no mundo digital geram cerca de 1000 *petabytes*³ por dia. De acordo com o *Statistica*

^{2, 3} Ver termos no glossário.

Digital Economy Compass, foram criados cerca de 33 *zettabytes*⁴ de dados apenas em 2018. Em termos físicos, tal equivale a utilizar 660 bilhões de discos *blu-ray*, 33 milhões de cérebros humanos, 330 milhões de discos externos de 100 *terabytes*⁵, 132 supercomputadores e 73 gramas de ADN.

No dia a dia, todos nós contribuímos para a criação de big data coletivo, aumentando o universo dos dados. Um exemplo disso é o *crowdsourcing*, uma prática cada vez mais popular, que consiste num processo de obter serviços, ideias ou conteúdos ao solicitar contributos da comunidade online utilizando a Internet (por exemplo, através de redes sociais). A Internet já não fornece apenas informação de forma passiva, mas alerta e guia ativamente os indivíduos para as melhores opções ao nível pessoal, enquanto os estimula a partilhar a sua experiência (como o *Tripadvisor* e o *Waze*). Os cidadãos começam a operar como sensores, criando dados e inteligência coletiva (a designada «Wisdom of crowds»), que podem ser utilizados pelas empresas para a sua diferenciação competitiva.

1.1.3. Big data é variedade

Além da explosão de grandes volumes de dados de forma significativa, existem vários tipos de big data. A definição de big data compreende dados estruturados, semiestruturados e não estruturados. Em geral, cerca de 20% do big data é estruturado (através, por exemplo, de tabelas e bases de dados de valores numéricos) enquanto 80% é semiestruturado e não estruturado (por exemplo,

^{4, 5} Ver termos no glossário.

registos da *web*, imagens, mensagens). Apesar de, até hoje, as ferramentas de *business intelligence*⁶ (BI) tradicional permitirem às empresas a gestão eficiente de dados estruturados, para a análise de informação não estruturada torna-se necessário recorrer a soluções de big data analytics.

Existem várias formas de criar dados: de pessoas para pessoas, de pessoas para máquinas e de máquinas para máquinas. Big data engloba todas estas formas, sendo o produto de transações, interações e observações.

Numa primeira instância, os sistemas de informação mais tradicionais (por exemplo, ERP⁷, CRM⁸) contêm informação muito estruturada e específica de transações feitas sobre operações de negócio (como o registo de compra). Ao entrar no ambiente *web*, já é possível observar informação menos estruturada, mais diversa e complexa de analisar (por exemplo, registos e acessos, histórico de ofertas, interações em redes sociais). Neste nível, são realizadas interações entre pessoas e máquinas. Por fim, quando estamos perante o universo de observações, existe um volume de informação muito maior, diverso e complexo. Estas observações tendem a vir de um tipo de big data específico: IoT (*Internet of Things*)⁹. IoT compreende diversos tipos de observações, desde sensores (objetos que medem vários fenómenos como temperatura, movimento, pressão, entre outros), conteúdos gerados pelos

⁶ Ver termo no glossário.

⁷ ERP – Enterprise Resource Planning – É um software assente num repositório único de informação que gere áreas críticas de uma empresa como finanças, cadeia de fornecimento, operações, comércio, relatórios, fabrico e recursos humanos. Permite ter uma visão integrada do negócio.

⁸ CRM – Customer Relationship Management – Os sistemas de CRM são sistemas de informação desenvolvidos com o objetivo de auxiliar na gestão do relacionamento com o cliente.

⁹ Ver termo no glossário.

utilizadores, mensagens. Esta última camada de dados não só é a mais densa de analisar, pela sua complexidade e variedade, mas também pelo volume.

1.1.4. Big data é veracidade

A veracidade de big data está relacionada com a qualidade e consequente confiança que se pode ter sobre os dados. Garantir a qualidade dos dados é fundamental para a análise eficiente de big data. A natureza dos dados pouco estruturados de big data faz com que exista muito ruído. Por este motivo, é fundamental que, associada a uma iniciativa de big data dentro das empresas, exista igualmente um processo de controlo de qualidade (por exemplo, regras de validação para exclusão de determinados dados).

1.2. ECOSSISTEMA DE BIG DATA E FERRAMENTAS

De forma a entender o ciclo de vida do big data e como funciona uma plataforma, torna-se relevante perceber o seu ecossistema. Tal como é possível visualizar na Figura 1.2, o ecossistema de big data é todo o fluxo de dados desde a sua criação até à sua utilização nos processos de negócio para suporte à decisão.

Assim, tal como é possível observar na figura, o ciclo de vida do big data inicia-se com a sua criação (1). Tal como já vimos anteriormente, o big data é criado por produtores de dados (desde pessoas, sistemas, sensores, etc.). Quando a organização decide começar a utilizar big data, deve ponderar investir numa componente

de aquisição dos dados (2). Existem várias ferramentas para obter dados, sendo *Hadoop*¹⁰ e *Cloud*¹¹ as mais comuns.

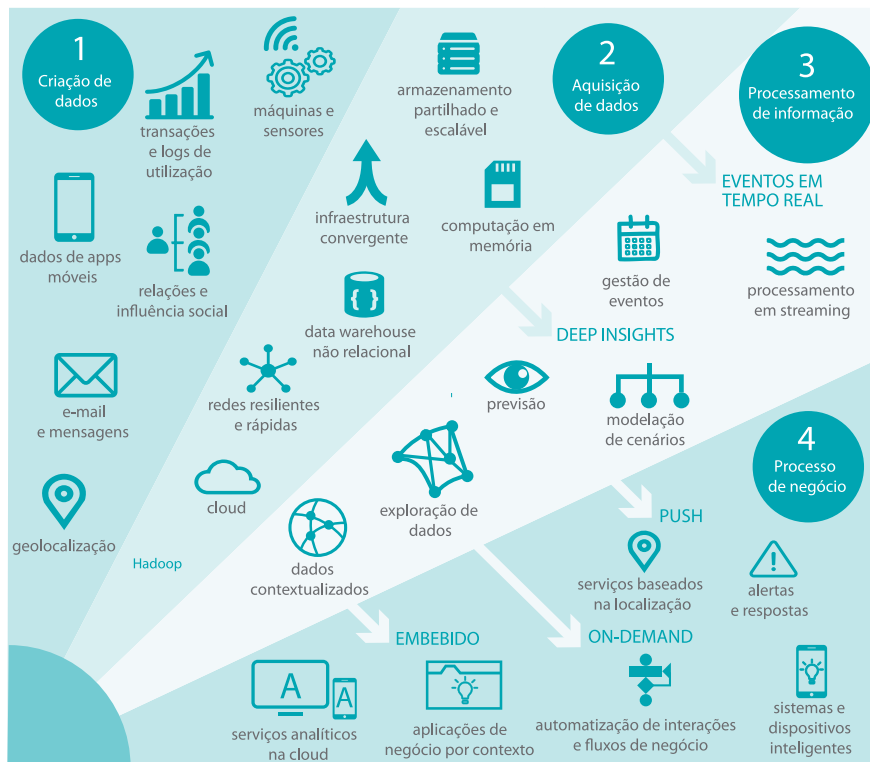


Figura 1.2 — Ecossistema de big data.
Fonte: IDC (adaptado).

Este ecossistema inclui a plataforma de big data, podendo ser composta por um conjunto de plataformas/ferramentas de aquisição de dados (*input*), processamento de informação (3) e inclusão dessa mesma informação nos processos de negócio (4) através de vários tipos de visualização (*output*).

^{10, 11} Ver termos no glossário.

TRANSFORME DADOS EM DECISÕES INTELIGENTES E POTENCIE A CRIAÇÃO DE VANTAGENS COMPETITIVAS.

Os dados são considerados o novo petróleo do século XXI e o verdadeiro ativo escondido nas empresas. Nunca se produziram, trocaram e compartilharam tantos dados ao atual ritmo exponencial e alucinante. A este fenômeno chamamos big data.

O aparecimento do big data e da *Internet of Things* veio desafiar a forma como guardamos, processamos, gerimos e analisamos esses mesmos dados, revolucionando a vida e o comportamento tanto da sociedade como das empresas. Nunca foi tão difícil selecionar informação fidedigna, pois são tantas as fontes de informação como de desinformação.

Nesta era digital, saber olhar para a informação certa no momento certo é crucial na tomada de decisão. Por isso, ser *data-driven*, deixar os dados falarem por si só e tomar decisões com base em *smart data*, é uma transformação profunda que, certamente, mudará o rumo dos negócios das empresas.

Neste guia prático, Nadine Côrte-Real, especialista em Gestão de Informação, concentra num só livro um conjunto de boas práticas e recomendações em várias vertentes da gestão de informação: governação, big data analytics e inteligência artificial responsável.




Através deste livro conseguirá obter conhecimento estratégico para:

- Gerir e governar informação de forma eficiente para o negócio, promovendo oportunidades de monetização dos dados e criando vantagens competitivas, seja qual for a indústria, a dimensão e contexto tecnológico da empresa;
- Potenciar a aceleração digital do negócio e criar um ecossistema de inovação e transformação digital alinhado com os objetivos estratégicos da empresa;
- Assegurar a adoção de ferramentas de governação, big data analytics e inteligência artificial para criar valor de negócio sustentável e contínuo;
- Promover o desenvolvimento de uma organização *data-driven* com foco na inteligência da informação e conhecimento do negócio.



Penguin
Random House
Grupo Editorial

Gestão

 penguinlivros.pt
  penguinlivros

ISBN 9789896235505



9 789896 235505 >