
V. 9, N. 2, ABR./JUN. 2019

Rafael Sena da Conceição
Universidade De São Paulo, SP,
Brasil
✉ rafaelsena@gmail.com

Francisco Carlos Paletta
Universidade De São Paulo, SP,
Brasil
✉ fcpaletta@usp.br

ARTIGO

Editor
Alfredo Passos
profdrpassos@gmail.com

RECEBIDO EM: 13/12/2019
APROVADO EM: 06/06/2019

© Atelier Brasil
Rua Pe. Guilherme Pompeu, nº1,
Centro- Santana de Parnaíba
06501-055 - São Paulo - Brasil

BIG DATA E A INFORMAÇÃO PÚBLICA - SUPORTE À TOMADA DE DECISÃO ESTRATÉGICA

BIG DATA AND PUBLIC INFORMATION – SUPPORT FOR STRATEGIC DECISION- MAKING

Resumo: O avanço da Tecnologia da Informação tem levado ao crescimento exponencial do volume de dados. O desafio que se apresenta é, diante deste universo, conseguir identificar, ao longo do processo de tomada de decisão estratégica, os dados com verdadeiro potencial de geração de valor à organização e, em seguida, transformar este potencial em vantagem competitiva. A metodologia desta pesquisa esta baseada em revisão da literatura e posterior estudo qualitativo e quantitativo, através da aplicação de questionário a provedores de informações de inteligência em Fazenda Pública. Os resultados demonstraram que, em regra, há um alinhamento de expectativas entre a alta administração do órgão público e os provedores de informação. A tomada de decisão estratégica e os resultados organizacionais são fortemente impactados pela gigantesca quantidade de dados disponíveis na era digital e a capacidade de sua organização e análise.

Palavras-chave: Big Data. Mineração de Dados. Inteligência Estratégica. Tomada de Decisão. Informação Pública. Informação Estratégica. Privacidade dos Dados.

Abstract: The advances in Information Technology have led to an exponential growth in the volume of data. The challenge that we face is that, as we approach this reality, we should be able to identify, in the process of strategic decision-making, the data with true potential for the generation of wealth for the organization and then turn this potential into a competitive advantage. The methodology of this study is based on a literature review followed by a qualitative and quantitative study, carried out through the application of a questionnaire to information intelligence providers in a tax administration organization. The results show that, as a rule, there is an alignment between the expectations of the high management team and those of the information providers. Strategic decision-making and the organizational results are heavily impacted by the huge amount of data available in the digital age and the capacity of organizing and analyzing them.

Keywords: Big Data. Data Mining. Strategic Intelligence. Decision-Making. Public Information. Strategic Information. Data Privacy.

I INTRODUÇÃO

Com o Big Data novos desafios associados à análise de dados apresentam-se com relevância estratégica no campo da Ciência de Dados e, adotando a célebre frase, atribuída a Peter Druck e W. Edwards Deming - “você não pode gerenciar o que você não pode medir” - como premissa fundamental, chega-se a conclusão que há um espaço significativo para o crescimento na importância da tomada de decisão de perfil racional. À medida que as ferramentas e filosofias associadas ao Big Data se espalham a maneira de se dirigir uma organização mudará de tal forma que uma revolução gerencial deve acontecer (MCAFEE e BRYNJOLFSSON, 2012).

A colocação acima é aderente à realidade dos órgãos públicos e em especial a uma Fazenda Pública¹, pois: i) diversas obrigações acessórias² são exigidas em formato eletrônico (ex.: Nota Fiscal eletrônica e declarações de bens de milhões de contribuintes), ii) pode fazer uso de informações oriundas de diversos bancos de dados públicos, iii) pode obter dados de outras Administrações Tributárias³. E considerando, ilustrativamente, que apenas a cidade de São Paulo tem hoje uma população estimada de aproximadamente 12 milhões de pessoas⁴, que acumulam algum tipo de renda, consomem e tomam serviços, é facilmente visualizável a gigantesca quantidade de dados disponíveis para os Fiscos. Ao mesmo tempo, os recursos são escassos, o que inviabiliza a fiscalização de todos os contribuintes e tornam imperioso, em prol da eficiência, que sejam priorizados aqueles que tenham o maior potencial arrecadatório. Neste momento, a área de inteligência, alinhada a estratégia, e tendo como insumo o universo de dados a sua disposição, pode contribuir de sobremaneira aos objetivos estratégicos.

Todavia, a operacionalização deste processo não é trivial: são necessários recursos humanos muito qualificados tanto em termos tecnológicos quanto em termos do modo de execução do negócio⁵ para conseguir extrair insights dos ativos informacionais da organização a partir da descoberta de relacionamentos dentro dos dados e na detecção de padrões (GARTNER). Também é necessária a existência de uma infraestrutura de TI que ofereça alta capacidade de processamento, pois os algoritmos empregados podem apresentar custos computacionais⁶ significativos e, com isso, consumirem um tempo de execução que comprometa a tempestividade do conhecimento proporcionado e prejudicando, por conseguinte, a tomada de decisão.

A despeito da relevância do tema em estudo, se não houver um efetivo processo comunicacional, de alinhamento top down da estratégia e um feedback por parte dos tomadores de decisão a respeito da

¹ No contexto deste trabalho, a expressão “Fazenda Pública” pode ser entendida como sinônimo de “Administração Tributária” e “Fisco”.

² São ações ou omissões que podem ser exigidas dos contribuintes a fim de prover informações relevantes à Administração Tributária. Previsão legal está contida no Código Tributário Nacional (CTN) em seu artigo 113, parágrafo segundo.

³ Previsão legal contida no CTN, em seu artigo 199, *caput* e parágrafo único.

⁴ Estimativa para o ano de 2018 elaborado pelo IBGE. Fonte: <https://cidades.ibge.gov.br/brasil/sp/sao-paulo/panorama>, acessado em 29/08/2018.

⁵ Estes recursos humanos são chamados de cientistas de dados e, usualmente, são dotados das seguintes características: dominam linguagens de programação, técnicas estatísticas avançadas e conhecem os processos de negócio da empresa. A coleção de tamanha vastidão de conhecimentos em um único ser humano naturalmente é rara, o que tornam estes profissionais carecidos no mercado de trabalho.

⁶ O custo computacional é uma métrica que indica quão rápida é a execução de um algoritmo e quantidade de memória utilizada. Sem maiores rigores técnicos, pode-se afirmar que se um algoritmo computacional tem menor custo, então ele exige um computador com menor capacidade de processamento para ser executado.

compreensão e utilidade das informações providas todo o esforço intelectual e financeiro pode restar desnecessário. Este estudo de caso busca verificar a conformidade a estes aspectos.

2 INFORMAÇÃO PÚBLICA

Batista (2010) conceitua informação pública como:

“Informação pública é um bem público, tangível ou intangível, com forma de expressão gráfica, sonora e/ou iconográfica, que consiste num patrimônio cultural de uso comum da sociedade e de propriedade das entidades/instituições públicas da administração centralizada, das autarquias e das fundações públicas. A informação pública pode ser produzida pela administração pública ou, simplesmente, estar em poder dela, sem o status de sigilo para que esteja disponível ao interesse público/coletivo da sociedade. Quando acessível à sociedade, a informação pública tem o poder de afetar elementos do ambiente, reconfigurando a estrutura social”.

Esta definição buscou delimitar bem o conceito de informação pública, mas ela pode ser vista como um conceito *stricto sensu*, pois atrela a ideia de publicidade ao Estado.

Alargando este conceito, é possível incluir como informação pública aquelas que são mantidas por entidades privadas, mas estão disponíveis ao acesso de todos. São exemplos cristalinos presentes apenas neste conceito *latu os* dados disponíveis em redes sociais.

Este novo conceito também torna ainda mais delicada a questão da privacidade e da intimidade, as quais são protegidas constitucionalmente⁷, pois os dados podem ser disponibilizados para o público sem a ciência do seu respectivo gerador⁸. Isto gera um alerta no processo de decisão: deve haver uma análise prévia a respeito da legalidade do uso da informação, mesmo que ela esteja disponível para todos⁹.

Em contraponto a informação pública existem as informações sigilosas, as quais podem estar em posse do Estado, mas não podem ser publicizadas em respeito aos comandos constitucionais presentes na Carta Magna de 1988, artigo 5º, inciso X, *in verbis*:

“São invioláveis a intimidade, a vida privada, a honra e a imagem das pessoas, assegurado o direito a indenização pelo dano material ou material decorrente de sua violação”.

A intimidade e a privacidade são conceitos considerados idênticos por muitos doutrinadores. Para outros, o direito à intimidade seria um subconjunto do direito à privacidade (MENDES e BRANCO, 2016). Estes autores conceituam o direito à privacidade, num sentido mais estrito, como:

Pretensão do indivíduo de não ser foco da observação por terceiros, de não ter os seus assuntos, informações pessoais e características particulares expostas a terceiros ou ao público em geral.

⁷ Previsão contida no art. 5º, inciso X da CF/88.

⁸ Um exemplo são os sites que coletam, organizam e compilam informações a respeito de uma determinada e nominada pessoa.

⁹ Visando obter a necessária conformidade jurídica é recomendável compreender as limitações contidas na Lei Geral de Proteção de Dados Pessoais (LGDP) – Lei no 13.709/2018.

Visando respeitar o comando constitucional são instituídos sigilos. No caso da administração tributária se trata do sigilo fiscal, o qual a obriga a não publicizar os dados dos contribuintes. O Código tributário Nacional, consoante a Constituição Federal, ordena no seu artigo 198, caput, in verbis:

“Sem prejuízo do disposto na legislação criminal, é vedada a divulgação, para qualquer fim, por parte da Fazenda Pública ou de seus funcionários, de qualquer informação, obtida em razão do ofício, sobre a situação econômica ou financeira dos sujeitos passivos ou de terceiros e sobre a natureza e o estado dos seus negócios ou atividades”.

E também, reforçando a relevância jurídica do tema, está presente no Código Penal, tipificado como crime, em seu artigo 325, in verbis:

“Revelar fato de que têm ciência em razão do cargo e que deva permanecer em segredo, ou facilitar-lhe a revelação: Pena – detenção, de 6 (seis) meses a 2 (dois) anos, ou multa se o fato não constitui crime mais grave”.

Percebe-se com isso, que i) um dado sigiloso não necessariamente é conhecido apenas pelo seu proprietário tal como uma análise mais superficial poderia levar a crer e ii) o sigilo fiscal é um ponto extremamente crítico, pois está no contexto de direitos fundamentais do ser humano.

A consequência prática é que análises e relatórios gerados com dados contendo sigilos não podem ser divulgados à sociedade civil e nem mesmo para outros órgãos, o que exige maior controle por parte dos provedores de informações.

3 INFORMAÇÃO ESTRATÉGICA

Informação estratégica é a informação que alimenta a administração estratégica de uma organização sob a ótica do planejamento, execução, controle e monitoramento; e revisão administrativa.

Estratégia, de acordo com Hitt, Ireland e Hoskisson (2011), pode ser definida como a coleção integrada e coordenada de comprometerimentos e ações para explorar as competências essenciais e ter retornos acima da média. Sinteticamente, é a definição das escolhas que uma organização realiza de acordo com estes autores. Caso os resultados propiciados pela estratégia sejam satisfatórios, ela se torna uma vantagem competitiva, segundo eles.

No contexto de uma Secretária da Fazenda, pode-se entender a vantagem competitiva como a capacidade da organização pública em implementar uma estratégia em que os contribuintes que não cumprem a legislação adequadamente possam ser identificados e autuados. Retornos acima da média podem ser compreendidos no contexto estudado como maiores valores de recuperação de receita¹⁰.

Desta conceitualização percebe-se a grande importância da informação estratégica na gestão de qualquer organização. Hitt, Ireland e Hoskisson (2011) ainda colocam que, na atual era da informação, a capacidade de

¹⁰ Recuperação de receita pode ser definida como a receita que o Estado tinha direito, mas que só foi realizada após uma ação do Fisco.

acessar e utilizar eficaz e eficientemente as informações se tornou uma fonte significativa de vantagem competitiva em praticamente todas as indústrias.

Estes autores discorrem:

“...conhecimento (informação, inteligência e expertise) é a base da tecnologia e da sua aplicação. No cenário competitivo do século XXI, o conhecimento é um recurso organizacional fundamental e está se tornando cada vez mais uma fonte valiosa de vantagem competitiva”.

E continuam:

“A probabilidade de obter competitividade estratégica no cenário competitivo do século XXI aumenta para a empresa que percebe que a sua sobrevivência está na capacidade de captar inteligência, transformá-la em conhecimento utilizável e difundi-la rapidamente por toda a empresa... Além disso, as empresas precisam criar rotinas que facilitem a disseminação dos conhecimentos locais por toda a organização para serem utilizados onde tiverem valor”.

A atuação de uma área de inteligência tal como previamente descrita está totalmente aderente às colocações acima postas.

4 BIG DATA

Big Data é composto de quatro dimensões: variedade, volume, velocidade e veracidade. Variedade se refere a diferentes formas de dados, estruturados ou não. Veracidade se refere a qualidade e acurácia dos dados. Velocidade trata do tempo utilizado no processamento, inclusive streaming. Volume se refere a escala dos dados, que podem alcançar facilmente terabytes (IBM, 2018)¹¹.

A variedade de dados advém da possibilidade de se lidar e mesclar dados estruturados (isto é, presentes em banco de dados relacionais), dados semiestruturados (exemplo: arquivos XML) e, principalmente, dados não estruturados (ex.: oriundos de redes sociais, áudios, vídeos, imagens, documentos de texto, conversas de chat, etc.). Estes últimos recebem destaque pela monta e também pela complexidade em serem alvo de análises. Enquanto os dados estruturados estão organizados a partir de classificações e relacionamentos de maneira que o processo de recuperação deles é facilitado, os dados não estruturados, por sua vez, têm baixíssima organização ou não têm nenhuma. Este segundo é um desafio tecnológico muito maior devido a sua magnitude¹² e complexo processo de armazenamento e recuperação. Como consequência desta complexidade, o processo de análise exige técnicas avançadas e maior capacidade de processamento para ser realizado.

Volume de dados é um ponto importante no contexto do Big Data porque caso ele seja pequeno as tecnologias existentes já são capazes de trata-lo. O problema surge quando estes volumes se tornam significativos (pode-se dizer que um volume é significativo quando atinge a ordem de grandeza de terabytes) de maneira a serem exigidas técnicas especiais para manipulação deles. É nesta realidade que o Big Data se insere.

¹¹ Esta definição é a que foi considerada a melhor no contexto específico deste trabalho, mas DAVENPORT (2014) apresenta considerações relevantes sobre os problemas das definições existentes de Big Data.

¹² Em 2012 existiam 2.8 zettabytes (2.8 trilhões de gigabytes) no mundo e apenas uma pequena parte deles é estruturada (DAVENPORT, 2014).

A velocidade é uma exigência do mundo dos negócios atual, onde o dinamismo impera, com mudanças ocorrendo permanentemente. Assim, é preciso tratar grandes volumes de dados de diferentes formas com uma velocidade que permite suportar à tomada de decisão de maneira tempestiva e acertada. Se não ocorrer desta maneira, finda-se o caráter utilitarista do Big Data.

Veracidade dos dados é premissa sine qua non: de nada adianta ter uma grande diversidade de dados e computadores poderosíssimos se os dados não estão corretos. Pois dados de entrada do processo de análise errados implicam, certamente, em resultados errados. A decisão fica desta forma, severamente comprometida ou mesmo totalmente inviabilizada. Por isso, a análise da qualidade dos dados deve ser uma das etapas iniciais de qualquer projeto de Big Data: sua falta é um importante indicador da inviabilidade técnica dele.

Não obstante, o ponto nevrálgico associado ao Big Data é realmente como analisa-lo (convertê-lo em insights, inovações e geração de valor para o negócio) de acordo com DAVENPORT (2014). São muitos dados existentes e boa parte dele não tem valor relevante para a organização. Identificar o que é interessante ou não para a organização pode ser um árduo trabalho que exige conhecimentos das regras de negócio e de métodos quantitativos, em especial, da mineração de dados. Trata-se de um processo de experimentação e descoberta desenvolvido pelas organizações que, por vezes, levam a resultados insatisfatórios. Neste contexto, o cientista de dados, graças a sua expertise, ganha papel de destaque.

Quando este processo é bem-sucedido os dados agregam tamanho valor à empresa que elas tendem a ser tornaram data-driven, isto é, elas passam a ter os seus processos orientados a dados.

5 MINERAÇÃO DE DADOS

Segundo Tufféry (2011), mineração de dados é um conjunto de métodos e técnicas para explorar e analisar conjunto de dados (que usualmente são grandes), numa maneira automática ou semiautomática visando encontrar nestes conjuntos regras desconhecidas ou escondidas, associações ou tendências. Em suma, ela busca extrair conhecimento dos dados. Percebe-se facilmente como a mineração de dados pode se relacionar bem com o Big Data: este oferece os dados necessários para a mineração de dados e esta permite a identificação e análise dos dados relevantes.

A Mineração de dados pode ser considerada um ramo da inteligência artificial e, de acordo com Tufféry (2011), pode ser descritiva e também preditiva: as técnicas descritivas (ou exploratórias) são projetadas para trazer à tona informações que estão presentes, mas ocultas pela massa de dados (ex.: no caso de clusterização automática de indivíduos e na busca por associações entre produtos ou remédios), enquanto as técnicas preditivas (ou explicativas) são projetadas para extrapolar as informações presentes em novas informações qualitativas (classificatórias ou de scoring) ou quantitativas (regressão).

Uma das principais ferramentas utilizadas na mineração é a estatística, que, segundo Business Dictionary (2018), é um ramo da matemática que se ocupa com a coleção, classificação, análise e interpretação de fatos numéricos para traçar inferências com base em probabilidades. Ela é dividida em descritiva e inferencial. A

primeira busca calcular quantidades que sumarizam e interpretam propriedades de uma amostra de dados. A segunda busca inferir propriedades de uma população a partir da análise de propriedades de uma amostra dela.

A mineração de dados é muito empregada no setor privado para suportar à tomada de decisão estratégica. Em relação ao setor público esta afirmativa também é verdadeira, como Ferreira e Negreiros (2008, p.3), citado por Ignácio (2010), pontuaram:

“Outra mudança, mais recente, da forma de atuação governamental que tem implicado novas demandas por informações foi a chamada focalização das políticas sociais. Várias das ações governamentais passaram a eleger segmentos específicos da população como seu público-alvo, requerendo para tanto, informações detalhadas das características da população, de modo a permitir a identificação do segmento prioritário, incluindo, em muitos casos, sua localização espacial e a construção de cadastros das famílias ou pessoas beneficiárias”.

Citando uma pesquisa de 2005, Tufféry (2011) coloca que o terceiro maior uso da mineração de dados é na detecção de fraudes (7% dos casos de uso). E este é um dos principais pontos atacados pelas Fazendas Públicas: a carga tributária é sempre um ônus para as pessoas físicas e jurídicas, afinal, trata-se de uma expropriação patrimonial. Para se furtarem desta obrigação, muitos contribuintes utilizam expedientes fraudulentos em documentos, declarações, sistemas e outros meios.

6 TOMADA DE DECISÃO

Segundo Bazerman e Moore (2014), a mente humana apresenta um poder superior ao dos computadores mais modernos, permitindo a realização de tarefas complexas. Entretanto, seus mecanismos de funcionamento não são bem compreendidos, impedindo, por exemplo, a antecipação de quando os processos cognitivos poderão falhar. Estes erros são causados, muitas vezes, por falhas de julgamentos.

Estes autores definem julgamento como os aspectos cognitivos do processo de tomada de decisão, composto das seguintes etapas, aqui apresentados sumariamente:

1. Definição do problema;
2. Identificação dos critérios relevantes (ex.: critérios morais, técnicos, legais, etc.);
3. Ponderação do peso de cada critério;
4. Geração de possíveis soluções do problema;
5. Identificação das alternativas que atendem cada critério
 - Esta etapa pode ser a mais complexa, pois pode exigir a predição do futuro: mapeamento das consequências potenciais de cada uma das possíveis soluções de acordo com cada critério identificado.
6. Identificação da solução ideal.

Bazerman e Moore (2014) afirmam que este processo apresentado acima é racional, lógico. Mas existem decisões tomadas de maneira rápida, automática, intuitiva. Para a tomada de decisão gerencial e, em especial, a estratégica, se recomenda o método racional, pois eles são de maior importância e este método permite um grau superior de reflexão.

Também colocam que o processo de decisão pode sofrer desvios em relação a racionalidade devido aos vieses e às heurísticas. Um exemplo de viés é o de confirmação, em que os dados são interpretados de maneira a suportar as conclusões que foram favorecidas no início. Heurísticas são estratégias simplificadoras que podem ser úteis no processo de decidibilidade, mas que se não bem estabelecidas, podem levar a deturpações no problema que levarão a decisões errôneas.

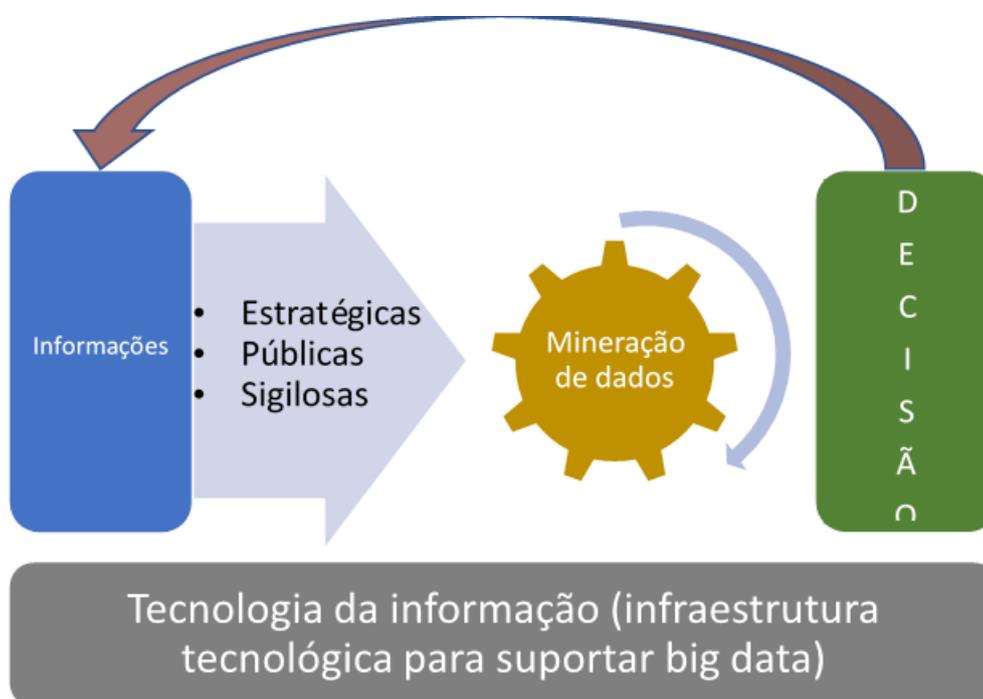
7 TECNOLOGIA, MINERAÇÃO DE DADOS E GESTÃO DA INFORMAÇÃO

As informações sejam elas estratégicas, públicas ou mesmo sigilosas são utilizadas como entradas para as técnicas de mineração de dados, que por sua vez, geram resultados que subsidiam o processo de tomada de decisão estratégica. A decisão tomada, direta ou indiretamente, gera novas informações de qualquer uma das três naturezas citadas, que retroalimentam o processo a fim de depurá-lo para alcançar melhores resultados.

A tecnologia acaba tendo, atualmente, papel relevante na administração. Ela permite que as organizações tenham a capacidade de lidar com grandes e diferentes problemas simultaneamente a menores custos, com maior rapidez e confiabilidade.

Os conceitos apresentados previamente trabalham, na realidade laboral, de maneira entrelaçada e coordenada. A imagem abaixo ilustra este processo:

Figura 1 - Fonte: elaborada pelos autores



Fonte: Elaborada pelos autores, 2019.

Tentando tornar mais tangível, a seguinte situação hipotética ilustra o explicado: dados de empresas são coletados em repositórios públicos conformes à legislação vigente (sendo assim, informações públicas), as quais são cruzadas com declarações tributárias (informações sigilosas, portanto), sendo priorizadas atividades econômicas de maior capacidade arrecadatória (informação estratégica). Esta miscelânea de informações alimenta as técnicas de mineração de dados que geram informações a respeito da probabilidade das empresas do setor econômico priorizado ser passíveis de inscrição na dívida ativa. Em seguida, são iniciadas fiscalizações que comprovam o previsto pela mineração. Como informações relativas a inscrições na dívida ativa são públicas¹³, gerou-se informações desta natureza. Por fim, estas novas informações são utilizadas com o fito de aprimorar a mineração de dados.

Todo este processo ocorre – e ainda mais: é viabilizado – pelo aparato disponibilizado pela tecnologia da informação (TI): hardware com processadores e memórias de alta capacidade e softwares capazes de executar operações complexas com relativa facilidade. Sem eles, a execução das técnicas de mineração de dados, repletas de complicadas operações matemáticas, seria impossível em larga escala.

O elemento mais significativo da TI no mundo dos negócios são os bancos de dados e sistemas de informações. Dentro deste último existem os Sistemas de Informações Gerenciais (SIG) que objetivam proporcionar informação para suportar a tomada de decisão e tem como base um banco de dados

¹³ Previsão legal contida no CTN, em seu artigo 198, §3º, inciso II do CTN.

(CHIAVENATO, 2004). Isso é decorrente do fato de que são eles que armazenam e manipulam os dados que alimentam as decisões das empresas.

Chiavenato (2004) ainda afirma que tecnologia tem criado uma nova filosofia para tornar a empresa mais produtiva, dinâmica e competitiva baseada não somente na organização do conhecimento da empresa, mas também em como visualizar e utilizar todas as informações internas e externas (ex.: públicas) a favor do negócio.

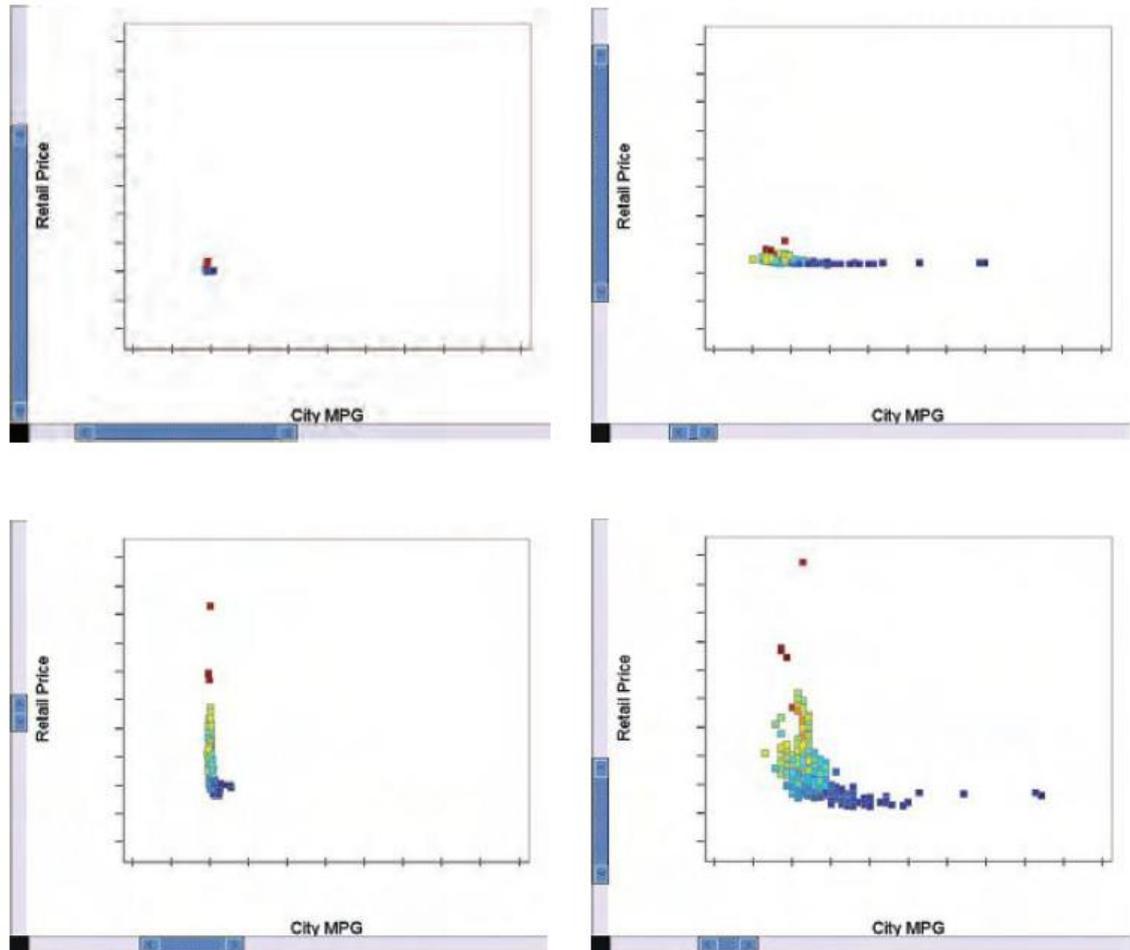
Neste diapasão, ele expõe que no processo de tomada de decisão a matemática e ciências correlatas, nas quais incluímos a mineração de dados, oferecem um tratamento lógico e racional através de uma abordagem quantitativa e determinística. O autor aponta como vantagens desta abordagem a otimização da execução de trabalhos e diminuição dos riscos envolvidos nos planos que afetam o futuro.

8 VISUALIZAÇÃO DE DADOS

Para Ward, Gristein e Keim (2010), abordagens visuais são um dos sentidos chave para compreensão de uma informação e, por conseguinte, para tomada de decisão. Esta relevância é relacionada à possibilidade de se evitar distorções nos dados e à capacidade de interpretação humana. É possível dizer, inclusive, que a forma como um dado é apresentado pode influenciar mais ou menos o tomador de decisão. Esta linha de pensamento corrobora com a colocação de Chiavenato (2004), já mencionada, de que a visualização é componente de uma nova filosofia para tornar a empresa mais eficiente.

Ward, Gristein e Keim (2010) apresentam um ilustrativo problema de distorção que a correta visualização de dados pode evitar. Trata-se do uso de escalas em gráficos: a mudança de uma escala pode levar a diferentes representações de um mesmo conjunto de pontos, como visualizável na imagem abaixo:

Figura 2



Fonte: WARD, GRINSTEIN e KEIM (2010)

Ward, Gristein e Keim (2010) ainda mencionam um estudo sobre como a forma de apresentação de dados escolhida leva a diferentes interpretações e, com isso, a diferentes decisões. Ele consistia em apresentar para diversos médicos os efeitos de um remédio nos pacientes em diferentes formas de apresentação gráfica (gráfico de barra, de pizza, tabela, etc.). A decisão a ser tomada era: interromper ou não o uso do remédio com base nos efeitos dele? Os pesquisadores perceberam que o acerto da decisão variou de 56% a 82% em função da representação gráfica exibida para o profissional da saúde. Vale destacar que os autores do estudo informaram além do tipo de representação gráfica, opções pessoais também influenciaram no resultado.

Com a explosão do volume de dados causado pelo Big Data, a necessidade de se ter ferramentas capazes de expor estes dados da melhor maneira possível a fim de facilitar e otimizar o processo de análise se torna ainda mais relevante.

Em complemento e reforçando a importância da visualização de dados na tarefa de facilitar a compreensão dos dados deve-se notar que as técnicas de mineração de dados e a estatística envolvem complexos cálculos que são de difícil compreensão para os não iniciados nos temas.

9 ESTUDO DE CASO

A tomada de decisão depende fundamentalmente de dados que a suportem com qualidade. Se os dados não têm qualidade, a decisão tomada também não terá (ANGELONI, 2003). Todavia, a qualidade dos dados é necessária, mas não condição suficiente. Outra variável que influencia na tomada de decisão é a interpretação destes dados: o dado pode estar correto, mas se o intérprete não apropria corretamente a informação associada ele, então, fatalmente, a decisão será prejudicada.

O estudo de caso em tela tentou aferir se os provedores de informações para a tomada de decisão estratégica no Fisco estão atentos a estas variáveis. Este grupo de pessoas é relacionada a área de inteligência, que tem sido valorizada nos últimos anos devido a sofisticação dos meios utilizados para a realização de ações ou omissões que visem o não pagamento dos tributos devidos e o avanço tecnológico, que viabilizou o processamento de grandes volumes de dados em uma janela temporal aceitável e com menores custos se comparado a quantidade de recursos humanos necessários para fazer esforço equivalente.

O conjunto de perguntas objetivou compreender se existe uma preocupação dos provedores de informações com o real valor e correto uso estratégico delas e suas razões. Para tanto, as perguntas foram agrupadas em dois grupos. O primeiro grupo mira tentar descobrir o perfil dos provedores de informações estratégicas: idade, grau e natureza da formação acadêmica. Este grupo está associado à identificação das razões da preocupação. O segundo grupo mira descobrir a mencionada preocupação e, se sim, qual o seu grau.

9.1 Resultados

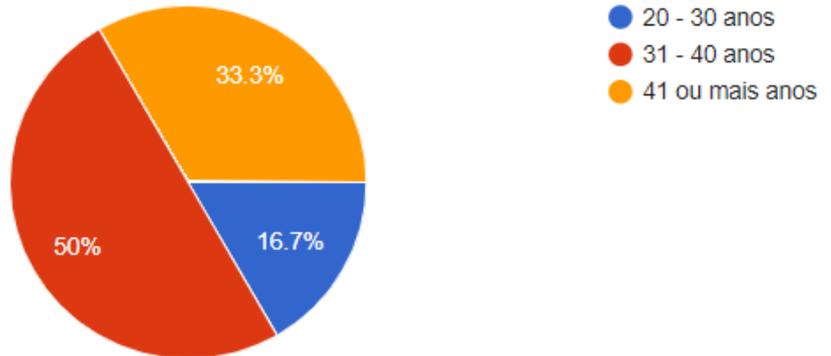
O questionário foi enviado para um pequeno grupo de pessoas (seis), pois a população alvo deste estudo, pelas suas próprias características, a saber: i) conhecimentos horizontal e verticalmente profundos e ii) detentora de acessos a dados sigilosos; é pequena.

Todos os convidados responderam ao questionamento, donde se levantam duas hipóteses para este grau de participação: ou o tema é de interesse de todos e, portanto, é realmente relevante ou é explicada por um fator psicológico, dado o constante apelo realizado pelos autores para que o questionário fosse respondido. O primeiro conjunto de respostas, que trata do perfil dos provedores de informações estratégicas, é composto pelas perguntas de número 1 a 4. Demais questionamentos compõem o grupo relacionando o grau de preocupação dos provedores de informações com o real valor e correto uso estratégico delas.

Figura 3

1. Qual é a sua idade?

6 responses



Fonte: elaborado pelos autores.

Tabela I

2. Qual é o seu curso de graduação?

6 responses

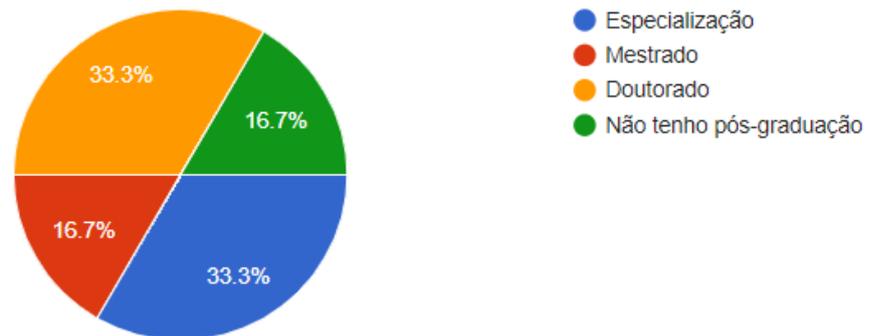
Comunicação social
Gestão em TI
Ciência da computação
Bacharelado em Ciência da Computação
Engenharia Civil
Engenharia Mecânica

Fonte: elaborada pelos autores

Figura 4

3. Possui pós-graduação? Informar o maior nível apenas.

6 responses

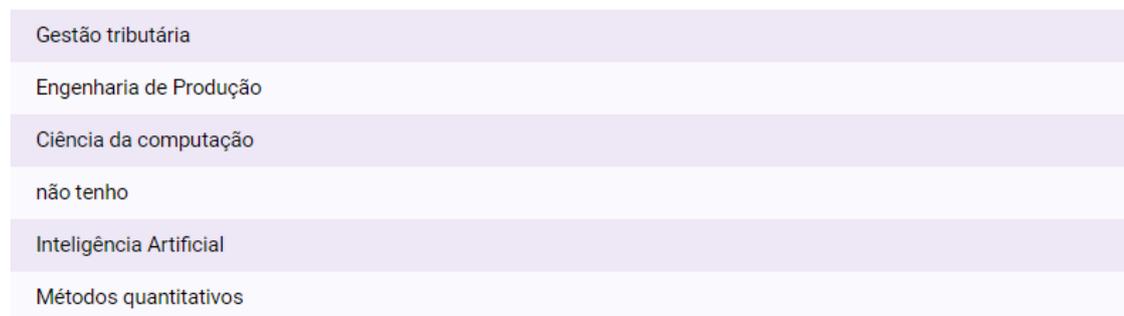


Fonte: elaborada pelos autores

Figura 5

4. Caso tenha pós-graduação, em qual área do conhecimento?

6 responses



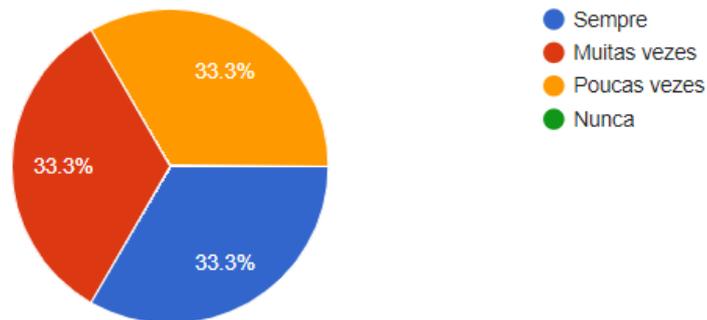
Fonte: elaborada pelos autores

Observa-se que o grupo de provedores tem perfil bastante heterogêneo: composto de pessoas de diferentes idades. A diversidade é importante, pois permite que soluções criativas sejam alcançadas através experiências diferentes, formações culturais e acadêmicas dessemelhantes. Entretanto, torna-se inviável traçar um perfil bem definido dos profissionais que estão atentos ao valor real e correto uso estratégico das informações por eles providas.

Figura 6

5. Você utilizada com que frequência técnicas estatísticas (básicas ou avançadas) para apresentar as informações?

6 responses



Fonte: elaborada pelos autores

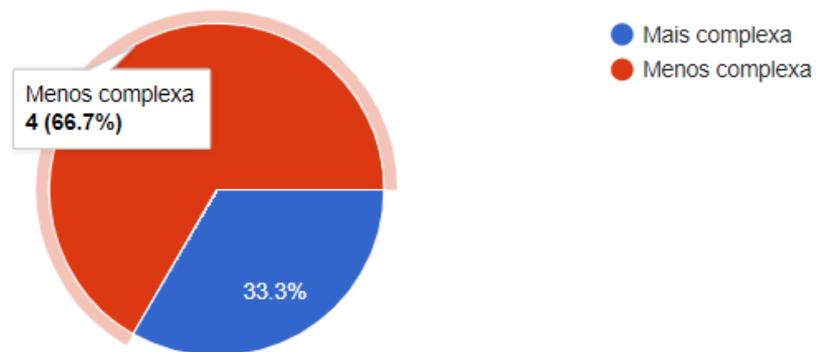
A estatística tem o poder de revelar padrões imperceptíveis a capacidade analítica humana e também tem capacidades preditivas dentro de certos limites. Mas a compreensão dos métodos utilizados é usualmente difícil e a não compreensão deles pode levar a errôneas interpretações dos resultados. Um exemplo: o usuário da informação tem ciência de que o método aplicado pode ter uma determinada probabilidade de erro? Se sim, sabe se ela é alta ou baixa? Mais concretamente: um método preditivo pode afirmar que certo evento vai acontecer com 99% de certeza ou com 55%. O tomador de decisão pode caminhar num sentido se souber que a certeza é de 99% e ir numa direção exatamente oposta se souber que ela é de 55%.

O exposto acima tenta evidenciar quão relevante é que o processo de comunicação das informações estatísticas seja realizado perfeitamente, a fim de evitar má compreensão. No caso estudado, dada a intensidade do uso de informações desta natureza, a questão se torna ainda mais relevante.

Figura 7

6. Entre uma forma de representação gráfica que seja simples, mas traga menos informação, e uma forma mais complexa, mas que agregue mais informações; qual você prefere?

6 responses



Fonte: elaborada pelos autores

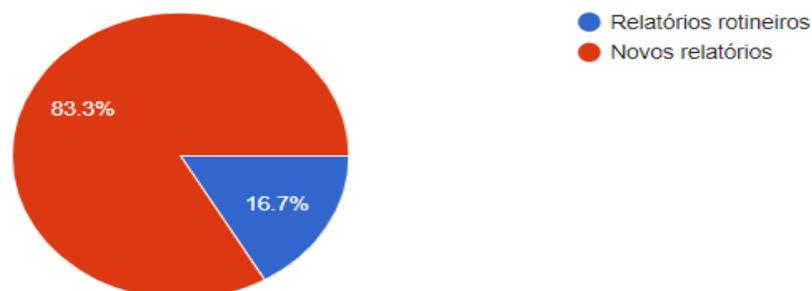
Uma tendência de profissionais altamente qualificados, como os questionados, é que eles busquem aplicar e demonstrar os conhecimentos avançados que possuem. Além disso, pode haver a crença de que quanto mais informação o gestor tiver, melhor será a decisão dele. Isto pode ser verdade, mas não é um dogma, simplesmente porque o volume de dados apresentado pode ser tão grande que a extração de uma informação simples pode se tornar algo complexo ou até mesmo levar a obtenção de uma informação errônea.

Pelo resultado da pesquisa observa-se que parte significativa dos profissionais está sensível ao tema e se preocupam em apresentar suas análises de maneira mais facilmente compreensível para o consumidor dos dados.

Figura 8

7. Das informações que você disponibiliza, há maior frequência em relatórios rotineiros ou em novos relatórios?

6 responses



Fonte: elaborada pelos autores

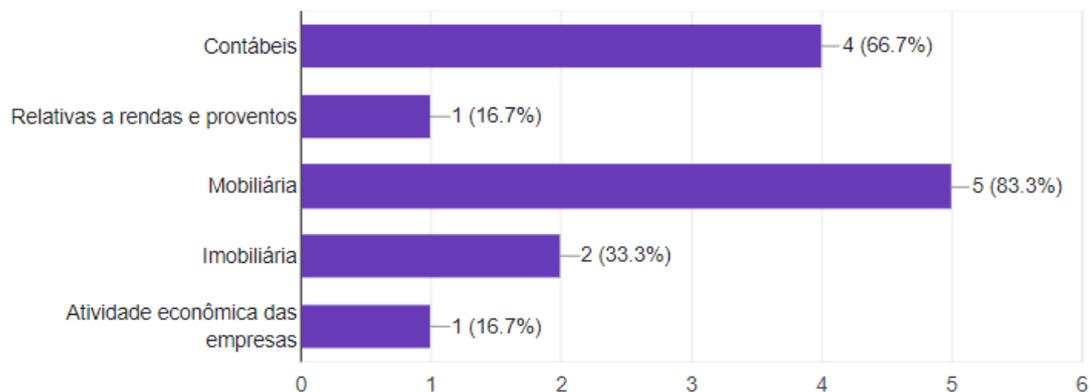
Há uma predominância de novos relatórios, o que é um resultado esperado visto que se trata de uma área de inteligência e, portanto, tem como um dos seus objetivos precípuos a obtenção e descoberta de novas informações.

Relatórios rotineiros podem ser gerados pela área por motivos de confidencialidade e/ou sensibilidade das informações, que exigem uma restrição de acesso a eles e que são relativamente comuns numa área envolta pelo sigilo fiscal.

Figura 9

8. Quais são suas principais fontes de informação?

6 respostas



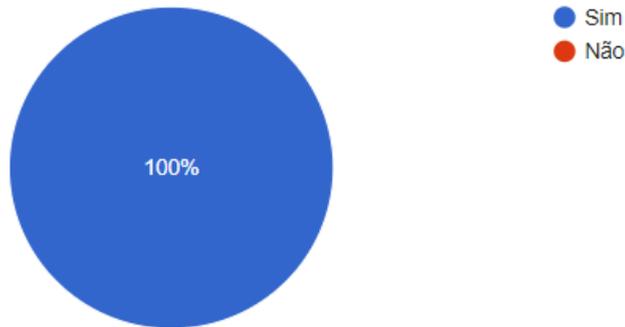
Fonte: elaborada pelos autores

As informações utilizadas pelas áreas de inteligência podem ser sigilosas ou públicas. Podem ser obtidas diretamente, via obrigações acessórias, ou indiretamente, via convênios ou leis com outros órgãos públicos ou ainda via entidades privadas, como Instituições Financeiras. Uma provável razão pela predominância de dados mobiliários é que esta área pode dar mais margens a fraudes do que as outras.

Figura 10

9. Você acha necessário se preocupar com o sigilo das informações?

6 responses



Fonte: elaborada pelos autores

Tabela 2

Por quê?

6 responses

Sigilo fiscal e lei de dados
Privacidade. Competitividade. Estímulo a inovação e reinvenção
Privacidade
O sigilo fiscal é protegido pela Lei 5.172/66 (CTN)
Privacidade, segurança e manutenção da competitividade de forma justa e leal
Trás segurança nas atividades

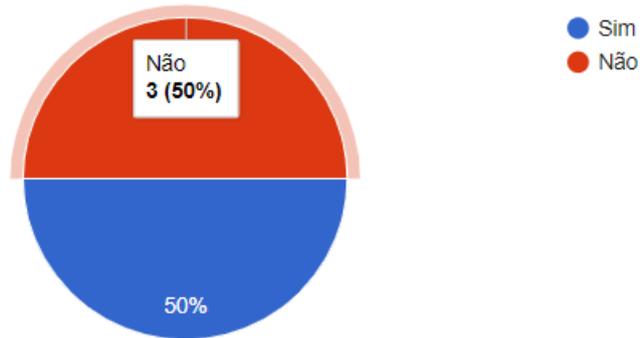
Fonte: elaborada pelos autores

Todos questionados estão atentos a necessidade de manutenção do sigilo das informações por variados e pertinentes motivos.

Figura 11

10. Você tem o hábito de coletar feedback a respeito da compreensão dos usuários das análises gráficas?

6 responses



Fonte: elaborada pelos autores

Tabela 3

Por quê?

4 responses

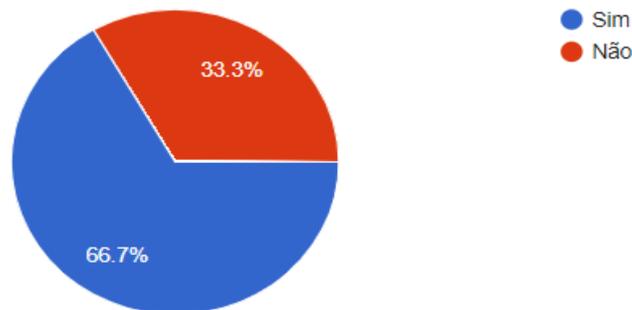
Para não parecer que estou achando que não foram capazes de entender.
Para ver se a informação que estou passando está sendo bem entendida.
Aprimorar os futuros gráficos e relatórios
É imprescindível para aprimorar a qualidade do trabalho

Fonte: elaborada pelos autores

Figura 12

11. Você tem o hábito de coletar feedback a respeito da utilidade das suas análises para a tomada de decisão estratégica?

6 responses



Fonte: elaborada pelos autores

Tabela 4

Por quê?

4 responses

Para manter-me alinhado à estratégia
Encaminho as análises aos superiores que costumam fazer comentários.
Avaliar a utilidade e aplicabilidade das análises
Permite direcionar as informações para aquilo que realmente importa para o administrador.

Fonte: elaborada pelos autores

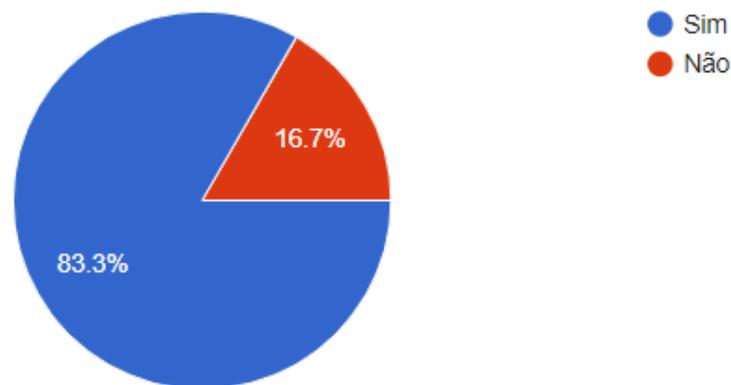
O feedback é importante porque permite ajustes objetivando um maior alinhamento estratégico e de expectativas entre a área de inteligência e os tomadores de decisão e, por conseguinte, colabora na obtenção de melhores resultados para a organização. De modo geral, as justificativas apresentadas na pergunta 10 e na 11 demonstram que os recursos humanos envolvidos têm consciência da importância do feedback.

Em que pese quem nem todos justificaram as respostas, é compreensível o posicionamento do profissional que não solicita o feedback. Considerando o relevo deste mecanismo, um caminho seria buscar formas de obtê-lo sem passar a impressão de incompetência do destinatário da informação. Uma solução seria institucionalizar este processo de feedback de tal maneira que o receptor fosse obrigado a fornecê-lo.

Figura 13

12. Você gostaria de receber os resultados desta pesquisa?

6 responses



Fonte: elaborada pelos autores

O resultado demonstra que boa parte dos questionados têm interesse em obter informações de caráter científico que podem promover alguma reflexão na forma de trabalho atual e, eventualmente, modificá-la. Isto combina com o perfil esperado de membros de um grupo de inteligência: investigativo, curioso e aberto a novas informações.

10 CONCLUSÃO

A informação estratégica tem grande valor para uma organização, seja ela pública, privada ou sigilosa. Ela é capaz de alterar os rumos da organização de maneira irreversível e é necessário que seus provedores tenham consciência disso e realizem seu trabalho com o máximo cuidado possível a fim de mitigar os riscos de que o caminho escolhido não seja o mais adequado.

De modo geral, pelos resultados analisados no questionário, os profissionais envolvidos estão cientes da importante e delicada posição em que atuam e, por isso, trabalham com a necessária diligência. Estas características evidenciam uma mão de obra qualificada, tal como se espera se agentes envolvidos num setor de inteligência. Provavelmente este quadro leva a resultados positivos para a organização estudada. Na análise do questionário hipóteses foram levantadas que para serem confirmadas exigem um aprofundamento das perguntas realizadas. Este pode ser um trabalho futuro.

REFERÊNCIAS

ANGELONI, M.. Elementos intervenientes na tomada de decisão. **Ciência da Informação**, v. 32, n. 1, maio 2003. ISSN ISSN 1518-8353.

BATISTA, Carmem Lúcia. **Informação pública** – entre o acesso e a apropriação social. 2010. Dissertação (Mestrado em Ciência da Informação) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo.

BAZERMAN, M. H.; MOORE, D. A. **Processo Decisório**. Tradução de Daniel Vieira. 8a. ed. Rio de Janeiro: Elsevier, 2014.

BD -BUSINESS DICTIONARY. **Business Dictionary**, 2018. Disponível em: <http://www.businessdictionary.com/definition/inferential-statistics.html>. Acesso em: 7 nov. 2018.

BD -BUSINESS DICTIONARY. **BusinessDictionary**, 2018. Disponível em: <http://www.businessdictionary.com/definition/statistics.html>. Acesso em: 7 nov. 2018.

BD - BUSINESS DICTIONARY. **BusinessDictionary**, 2018. Disponível em: <http://www.businessdictionary.com/definition/descriptive-statistics.html>. Acesso em: 7 nov. 2018.

CHIAVENATO, I. **Introdução à teoria geral da administração**. 7. ed. Rio de Janeiro: Elsevier, 2004.

DAVENPORT, T. H. **Big data @ work: Dispelling the Myths, Uncovering the Opportunities**. Boston: Harvard Business Review, 2014.

GARTNER. Gartner IT Glossary. **Gartner**. Disponível em: <https://www.gartner.com/it-glossary/data-scientist>. Acesso em: 29 set. 2018.

HITT, M. ; IRELAND, R. ; HOSKISSON, R.. **Administração estratégica: competitividade e globalização**. 2. ed. São Paulo: Cengage Learning, 2011.

IBM. **IBM Big Data Hub, 2018**. Disponível em: <https://www.ibmbigdatahub.com/infographic/four-vs-big-data>. Acesso em: 17 nov. 2018.

IGNÁCIO, S. A. Importância da estatística para o processo. **Revista Paranaense De Desenvolvimento**, jan./jun. p. 175-192, 2010.

MCAFEE, A.; BRYNJOLFSSON, E. Big Data: the management revolution. **Harvard Business Review**, p. 1-9, 2012.

MENDES, G. F.; BRANCO, P. G. G. **Curso de direito constitucional**. 11a. ed. São Paulo: Saraiva, 2016.

TUFFÉRY, S. **Data mining and statistics for decision making**. [S.l.]: Wiley, 2011.

WARD, M.; GRINSTEIN, G. G.; KEIM, D. **Interactive data visualization foundations, techniques, and applications**. Boca Raton: CRC Press, 2010.