

Estágios de Adequação Organizacional Sensível às Configurações *Data-Driven*

Alencar Santos Buriti Junior¹, Pedro Rafael Simplicio Cunha¹, Breno Rodrigo da Silva di Pippo¹, Eliomar Araújo de Lima¹

¹Instituto de Informática - Universidade Federal de Goiás - Goiânia, GO - Brasil

alencarburitijr@gmail.com, pedrorafael.sistemas@gmail.com,
brenouro@gmail.com, eliomar.lima@ufg.br

Abstract. *Organizations deal with their information artifacts in different ways and have different organizational configurations. A guiding axis can provide subsidies for solving problems and making business decisions. The methodological path consisted of an exploratory study, embodied in the form of a synthetic narrative review. The research focused on the discussion of different data-sensitive configuration models, with a focus on becoming a data-driven organization, supported by the conceptual approach and the historical context. As a result, it was possible to propose hypotheses relating organizational configuration models and analytical approaches with enabling factors for a data-driven journey. The study also made it possible to qualify the approaches for developing exploratory data analysis, including the CRISP-DM, for presenting an evolutionary and incremental cycle that favors its application.*

Resumo. *As organizações lidam com os seus artefatos de informação de diferentes modos e possuem configurações organizacionais distintas. Um eixo norteador poderá prover subsídios para a resolução de problemas e tomada de decisão de negócio. O percurso metodológico consistiu em um estudo exploratório, consubstanciado na forma de uma revisão narrativa sintética. A pesquisa centrou-se na discussão de diferentes modelos de configuração sensíveis a dados, com foco em se tornar uma organização data-driven, amparada pelo enfoque conceitual e do contexto histórico. Como resultado, foi possível a proposição de hipóteses relacionando os modelos de configuração organizacional e abordagens analíticas com fatores habilitadores para uma jornada data-driven. O estudo também possibilitou qualificar as abordagens para desenvolvimento de análise exploratória de dados, dentre os quais o CRISP-DM, por apresentar um ciclo evolutivo e incremental que favorece a sua aplicação.*

1. Introdução

Em geral, os setores da economia possuem como um dos elementos críticos para manter os níveis de competitividade e produtividade satisfatórios a capacidade de geração, produção e aplicação eficiente do conhecimento gerado pela combinação de informações e artefatos produzidos a partir do processamento de dados massivos. A competitividade utilizando dados como diferencial é vista não apenas como um mero

trabalho a ser feito, mas como uma decisão estratégica vital para a organização e necessária para que ela continue evoluindo.

Em diferentes áreas e segmentos, desde o atendimento ao cliente, experiência de compra, soluções e produtos, custos e risco, passando pela maximização de oportunidades, novos negócios até a tomada de decisão efetiva, são alguns exemplos que podem ilustrar o quanto os conglomerados de dados podem passar a se tornar um diferencial. Compreender que o processo de decisão orientado por dados processados massivamente não é feito apenas por tecnologias, é um passo fundamental em direção ao sucesso desse conceito.

A falta de equipes multidisciplinares, executivos que não compreendem a importância da utilização dos dados e uma cultura organizacional que não apoia totalmente a orientação por dados podem ser considerados como alguns dos pontos de fracasso em implementações do conceito *data-driven* (sensível a dados). Não se pode assegurar que o investimento financeiro por si só garanta que uma organização conseguirá ter êxito na adoção de um modelo orientado a dados. É necessário ter uma visão clara do caminho que é necessário perfazer, caso a organização entenda que o processamento de dados massivos são um diferencial estratégico.

Mapear o que a organização já consegue fazer de útil com os dados processados é o primeiro passo para entender quais serão os próximos e o que deve ser aprimorado. O interesse que as organizações têm em utilizar cada vez mais dados para se manterem competitivas e responder rapidamente às mudanças de mercado traz consigo uma enorme quantidade de possibilidades de implementação. Em meio às possíveis alternativas, uma pode ser pela manutenção do atual estágio de desenvolvimento, seja porque falta clareza do caminho a seguir seja porque a organização não possui capacidade de realização, impedindo a evolução analítica supostamente esperada.

O principal objetivo desta pesquisa é identificar fatores habilitadores para definição de uma configuração organizacional sensível a dados. Os fatores habilitadores são elementos constituintes para que uma organização possa evoluir analiticamente. A utilização desses fatores contribui para a jornada *Data-Driven* ao passo que cada configuração organizacional, metodologias e estágios de evolução indicam um possível caminho a ser seguido para essa evolução.

Para atingir o objetivo proposto, buscou-se identificar quais os pressupostos para que a mudança na configuração organizacional pudesse evoluir analiticamente. Adicionalmente, buscou-se entender quais habilidades são necessárias na construção de uma organização que utiliza a ciência de dados para se tornar mais competitiva, bem como definir quais os componentes organizacionais devem ser mapeados para descrever o estágio atual de seu modelo analítico e como evoluir a partir disso.

Este estudo busca evidenciar e identificar, por meio de uma breve revisão da literatura, de caráter exploratório, potenciais fatores habilitadores e propor hipóteses que implicam na jornada para uma organização escalar modelos sensíveis a dados, diretrizes para que organizações possam se preparar do ponto de vista de design organizacional para atingir o modelo *data-driven*.

A Seção 1 apresenta e situa o leitor sobre o objetivo do artigo, assim como faz uma breve contextualização. A Seção 2 caracteriza o conceito *data-driven* e aborda formas de configuração organizacional que contribuem para que uma organização evolua analiticamente. A Seção 3 aborda os conhecimentos necessários para que projetos analíticos possam evoluir na organização e também apresenta a metodologia de mineração de dados CRISP-DM. A Seção 4 apresenta a orientação metodológica deste estudo, para em seguida discorrer sobre os estágios evolucionários por meio dos quais uma organização pode ascender em busca de uma orientação sensível aos dados. É abordado também a análise e discussão dos resultados que foram obtidos durante a pesquisa. A Seção 6 traz as considerações finais da pesquisa.

2. Configurações Organizacionais Sensíveis a Dados

As estruturas organizacionais não são elaboradas apenas por organogramas ou documentos que apresentam os aspectos funcionais e de controle, mas de artefatos que podem revelar o posicionamento e o modelo de negócio para propiciar a estratégia da organização [Matias 2018].

A evolução das organizações com objetivo de alcançar um alto nível de competitividade tem resultado em uma nova configuração organizacional com uma estrutura adequada para melhor controlar os indicadores de desempenho, obter e analisar diferentes dados existentes e identificar recursos de acordo com a necessidade [Camargo 2018].

Uma mudança possível na estrutura organizacional é a criação de um departamento dedicado à análise de dados. Pode ser uma unidade combinada ou substituída por uma unidade de análise quantitativa, ou uma unidade completamente nova. Algumas grandes organizações possuem unidades ou departamentos dedicados à análise de dados completamente segregados da área de apoio à decisão [Sharda, Delen and Turban 2019].

Para [Griffin and Davenport 2014], quando se pensa em implementar um ecossistema de análise, é importante observar a cultura e visão da organização, para que o modelo de análise escolhido satisfaça os objetivos dessa organização. Pessoas e configurações organizacionais têm grande relevância para que os dados sejam utilizados orientando as decisões da organização.

2.1 Organização *Data-Driven*

Quando uma organização passa a assumir o conceito *data-driven*, ela passa a entender que uma pergunta ou um problema na organização será respondido via processamento de dados massivos e os artefatos de informação e informações reveladas poderão subsidiar a decisão a ser tomada [Vilvert 2020].

Podemos considerar que uma organização é *data-driven* quando sua tomada de decisão é baseada totalmente em dados processados e, diferentemente de uma abordagem *data-informed* (informada por dados), não há a utilização de intuição por parte dos tomadores de decisão [Stewart 2019].

Se por um lado a tomada de decisão *data-driven* propicia decisões sólidas e bem-informadas, por outro pode levar a decisões precipitadas, errôneas e causar prejuízos, caso não haja dados suficientes ou não se tenha uma equipe bem qualificada para delinear o modelo analítico [Kampakis 2020].

Além da importância de se ter dados de qualidade, [Teixeira 2018] descreve 6 pilares para que as organizações possam sustentar um modelo de operação sensível a dados e que almeja o crescimento exponencial por meio do investimento em tecnologia, equipes, processos e conhecimento dos clientes:

1. organizar dados de diversas fontes de forma rápida, ágil e padronizada;
2. obter os dados unificados dos clientes;
3. utilizar inteligência artificial, principalmente modelos preditivos para buscar respostas sobre tendências e visão futura;
4. treinar e ter times que sejam multidisciplinares e autônomos;
5. trabalhar com tecnologias que permitam fazer a maioria das coisas em tempo real; e
6. respeitar a privacidade das informações.

2.2 Configuração de Organizações Analíticas

De acordo com [Griffin and Davenport 2014], nenhum conjunto de estruturas e processos organizacionais é permanente, a liderança da organização deve estar disposta a mudar o modelo de análise se for necessário, para que ele continue atendendo as necessidades do negócio.

As configurações analíticas de uma organização podem compreender o modelo centralizado, no qual os analistas se posicionam em um centro de excelência; o modelo descentralizado, no qual a análise é incorporada nas unidades de negócio de maneira individual e; o modelo híbrido, que busca combinar uma unidade centralizada com áreas analíticas integradas às unidades de negócio [Miranda 2018].

O modelo centralizado atua como uma unidade principal e atende toda empresa. Utilizado, geralmente, quando há uma necessidade premente de análise amparada por dados na organização. O ponto forte dessa configuração consiste na capacidade de mostrar para o alto escalão o valor das análises abrangendo as áreas funcionais da empresa, até o nível de projeto. Seu ponto fraco decorre do fato de que cada unidade de negócio atendida por esse modelo pode tender a analisar um problema que não seja estratégico e, assim, desperdiçar esforços.

Segundo [Griffin and Davenport 2014], para a maioria das organizações, grupos de análise centralizados são mais eficazes, principalmente quando o trabalho dos analistas é integrado com as unidades e funções do negócio. Independente do modelo escolhido, faz-se necessário reavaliar o modo periodicamente para que ele continue atendendo as necessidades do negócio.

3. Em direção à Aplicação em Ciência De Dados

Conceitua-se Ciência de Dados como uma filosofia de estudo cujo propósito é melhorar a tomada de decisão, gerar ações para apoiar a gestão através de *insight* e obter resultados aplicando processos em problemas específicos do mundo real. Um conjunto de princípios deve ser executado escolhendo os algoritmos que melhor se adequam às necessidades, extraindo padrões não óbvios de grandes conjuntos de dados [Kelleher and Tierney 2018].

[Rautenberg and Carmo 2019] apresentam 3 áreas, sendo elas: programação de computadores, matemática e estatística e conhecimento de domínio. Os subconjuntos dessas áreas, provenientes da associação entre elas, podem ser, a título de exemplo: a associação da área de Ciência da Computação com a área de Matemática e Estatística obtém-se a ramificação da Inteligência Artificial acompanhada do Aprendizado de Máquina. Cada área de conhecimento citada é importante no desenvolvimento de projetos de ciência de dados e no entendimento que é necessário montar times de analistas e cientistas de dados multidisciplinares.

De outro ponto de vista, na compreensão de [Kelleher and Tierney 2018], entende-se a Ciência de Dados como um conjunto de domínios que são interdependentes em termos de comunicação, domínio de perícia, ética de dados e regulação, ciência da computação, visualização de dados, estatística e probabilidade e aprendizado de máquina.

3.1 CRISP-DM

Dos processos e metodologias que existem atualmente, o *Cross Industry Standard Process for Data Mining* – CRISP-DM – é provavelmente o mais utilizado e conhecido. Essa orientação metodológica pode ser bastante interativa e deve ter uma atenção especial já nos passos iniciais para evitar a obtenção de conhecimento a partir de perguntas mal formuladas [Sharda, Delen and Turban 2019].

Muitas vezes, encontrar a solução de um problema é um processo repetitivo de descobertas. É importante entender que nem sempre os dados terão uma relação exata com o problema a ser resolvido. O fator crítico dessa fase é fazer uma estimativa de custo e benefício para cada fonte de dados e verificar se precisará fazer mais investimentos para adquirir mais conjuntos de dados ou se apenas os dados disponíveis são suficientes.

O CRISP-DM possui 6 fases às quais tem por objetivo compreender desde a necessidade do negócio até a implementação das respostas ou soluções para os problemas envolvendo dados. As fases compreendem: entendimento do negócio; entendimento dos dados; pré-processamento dos dados; modelagem; avaliação e; implementação.

É importante compreender que a primeira iteração desse ciclo pode não trazer os resultados esperados e isso não quer dizer que o projeto tenha fracassado. Após cada iteração, a equipe de ciência de dados consegue compreender muito mais sobre o problema e ajustar o caminho [Provost and Fawcett 2016].

4. Metodologia Da Pesquisa

A pesquisa é de natureza qualitativa e de cunho exploratório e teve como foco estabelecer uma trajetória preconizada pelo método de revisão narrativa, que consiste em apresentar uma discussão sobre uma questão ampla, apropriada para descrever e discutir o desenvolvimento do tema sob o ponto de vista conceitual e histórico.

O ponto de partida é a manifestação de interesse em mapear os fatores impulsionadores para mudanças organizacionais que visem a orientação por dados nas organizações. Foram levantadas diferentes configurações organizacionais de modo a comportar uma área de desenvolvimento de modelos analíticos ou de ciência de dados. Esse levantamento se fez necessário para entender quais modelos se destacam, além de poder mapear as configurações sensíveis a dados.

Sendo um trabalho de revisão, teve como foco principal a descoberta de fatores habilitadores para que as organizações possam ascender a modelos sensíveis a dados. Esses fatores foram também utilizados como norteadores, visto que na jornada para se tornar uma organização *data-driven* foram elencados estágios propostos por [Davenport and Harris 2018], como forma de identificação do atual nível da organização e quais seriam os requisitos para avançar por cada estágio posterior até atingir o nível de competidor analítico.

5. Jornada Para Tornar-se Uma Organização *Data-driven*

A partir da década de 2010, uma miscelânea de modelos de processos, métodos, ferramentas e metodologias vêm sendo apropriados para subsidiar abordagens de transformação digital que possam impulsionar uma organização em sua jornada de assimilação tecnológica, por vezes disruptivas.

A transformação digital tende a criar uma organização de alta performance ao buscar inovação. Com isso, novas formas de implementação da transformação da cultura organizacional, maneiras diferentes de realizar as atividades de negócio que são impactadas pela conjuntura tecnológica atual, são esperadas [Silva 2018].

O objetivo da jornada é dar suporte às organizações para tomar suas decisões, o que requer elementos que lhe permitam caracterizar o problema, compreender o ambiente que cerca as decisões e identificar os impactos inerentes no qual a alta direção pode tomar suas decisões com mais segurança [Morais 2000].

Para [Sharda, Delen and Turban 2019], utilizar um processo padronizado ou mesmo um passo a passo, aumentará muito a chance de sucesso de um projeto de mineração de dados e descoberta de conhecimento.

Na composição de equipes de trabalho deve se ater a determinadas competências: apropriar a lógica, estrutura de dados e ferramentas disponíveis no mercado. No entanto, é imprescindível que os profissionais que atuarão nos projetos de Ciência de Dados tenham os princípios éticos, da segurança e da proteção de dados no âmbito pessoal e corporativo [Kelleher and Tierney 2018].

Em se tratando dos líderes, estes necessitam ser capazes de instigar a capacidade criativa das equipes com o objetivo de criar condições favoráveis para estimular os integrantes a entenderem as problemáticas que emergem das desconexões e níveis de abrangência corporativa e de negócio, contribuindo para a construção de potenciais soluções integradas [Silva 2018]. A liderança analítica leva a organização a ter um diferencial competitivo e cognitivo.

No contexto tecnológico, a partir da descrição da arquitetura da informação que permite absorver a complexidade organizacional, tem-se que o grande desafio encontrado no próprio ambiente organizacional é conseguir integrar as informações geradas dentro do ambiente eletrônico/digital, em formato compreensível [Teixeira and Valentim 2012].

Para [Davenport and Harris 2018, p. 166], “mudar os processos de negócios e o comportamento dos colaboradores é sempre a parte mais difícil e que demanda mais tempo em qualquer grande mudança organizacional”, ensejando o desenvolvimento de competências analíticas de modo contínuo. No mesmo estudo, os autores asseveram que são necessários de 18 a 36 meses de trabalho regular com os dados para começar a desenvolver uma linha estável de *insights* que possam ser transpostos para a prática.

Além de um longo caminho a ser percorrido para obter resultados relevantes, é necessário observar que a maioria das organizações não possuem capacidades analíticas bem desenvolvidas e nem mesmo um plano para que essas capacidades possam ser impulsionadas. Dessa forma, os 5 estágios que serão apresentados descrevem os comportamentos, recursos e desafios típicos para um desenvolvimento analítico, conforme ilustrado na Figura 1 [Davenport and Harris 2018].

O estágio 1, denominado “Pré-requisitos para uma análise competitiva”, em que as organizações precisam primeiramente se preocupar com o seu ambiente de dados de transações, o qual requer dados consistentes, corretos e de qualidade para a tomada de decisão [Davenport and Harris 2018]. Nesse estágio, a organização deve compreender a capacidade analítica em 3 áreas:

- organizacionais: conhecimento dos vetores de desempenho; escolha de uma competência distintiva; gestão de desempenho e execução da estratégia; integração e redesenho do processo;
- humanas: liderança e comprometimento da alta administração; estabelecimento de uma cultura baseada em fatos; desenvolvimento e retenção de competências; gerenciamento do pessoal analítico; e
- tecnológicas: dados de qualidade; tecnologias analíticas.

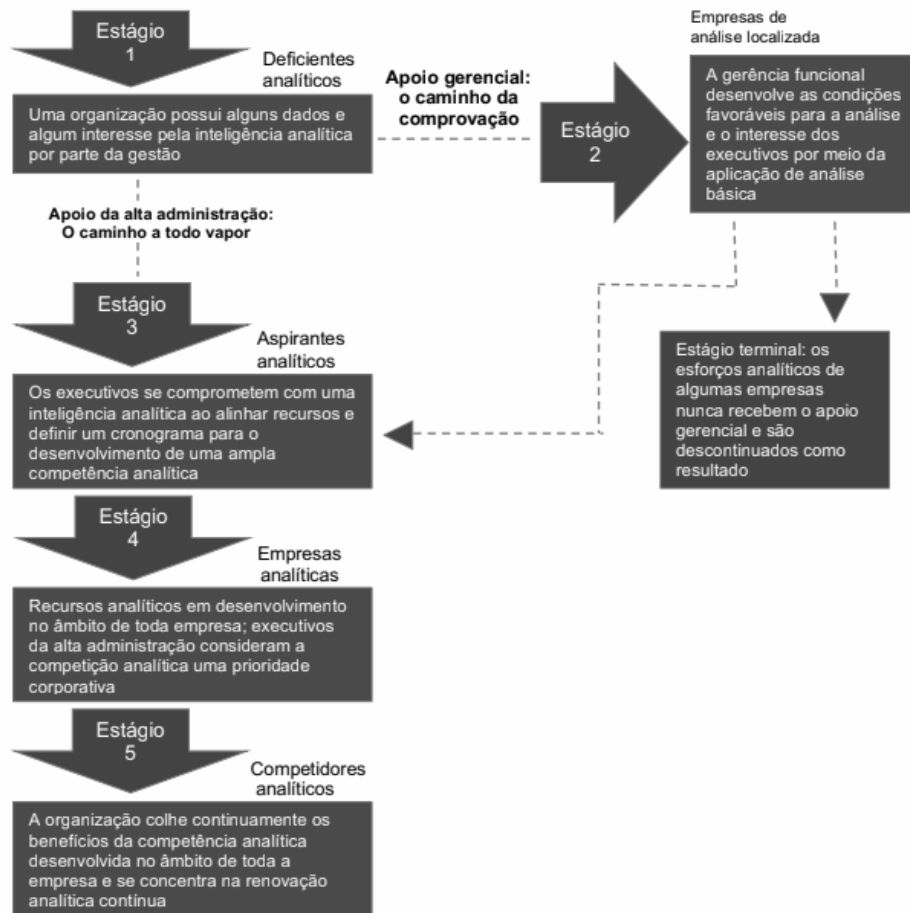


Figura 1. Roteiro para se tornar um competidor analítico
Fonte: Davenport; Harris: 2018, p. 167.

O estágio 2 é constituído pelo desvio de prova, compreendendo os executivos e organizações que já reconhecem os resultados obtidos através da competição analítica, ou inversamente, por meio da adoção de abordagem de teste e aprendizado que pode ocorrer através de análises em passos menores. Para cada percepção incremental do negócio é criado um impulso dentro da organização em favor da mudança para estágios mais elevados de competitividade analítica [Davenport and Harris 2018].

Uma organização pode permanecer no estágio 2 indefinidamente se a alta direção não conseguir perceber os resultados. De outro modo, poderá avançar para o próximo estágio, desde que [Davenport and Harris 2018, p. 175]: (1) encontre um patrocinador e um problema de negócios que pode se beneficiar da análise; (2) implemente um projeto pequeno e localizado para agregar valor e produzir benefícios mensuráveis; (3) documente os benefícios e compartilhar as notícias com as principais partes interessadas; (4) continue a construir uma série de sucessos localizados até que a organização tenha adquirido experiência e patrocínio suficientes para avançar para o próximo estágio.

O estágio 3 é constituído de aspirações analíticas e tem como principal característica o momento que a análise obtém um patrocinador executivo. Neste estágio,

é fundamental definir um conjunto de medidas de desempenho que sejam alcançáveis e processos devem ser implementados para monitorar o progresso. Algumas tarefas que devem ser executadas e exige análises mais sofisticadas como conhecimento analítico especializado juntamente com a adição de novas tecnologias analíticas [Davenport and Harris 2018].

O estágio 4 traz as competências analíticas desenvolvidas de forma madura em toda empresa. Nessa fase a organização obtém um progresso considerável na construção de patrocínio, cultura, habilidades, *insights* estratégicos, dados e tecnologia necessários para a competição analítica. Portanto, à medida que a organização aprende mais com cada análise, ela obtém um rico conjunto de novos *insights* para explorar e obter vantagem competitiva [Davenport and Harris 2018].

O estágio 5 é considerado o estágio final alcançado por organizações que se tornaram líderes analíticos. A análise agora é vista não apenas como um recurso importante para a organização, mas como a chave de sua estratégia e vantagem competitiva [Davenport and Harris 2018]. De acordo com os autores (*ibidem*, 2018, p. 185), “[...] a análise é utilizada para descrever inovações em toda a empresa”. Nesse estágio, o uso de modelos prescritivos e, principalmente, preditivos têm grande destaque e ajudam no crescimento dos produtos e serviços oferecidos pela empresa. Alcançar o estágio de líder analítico requer muito mais que entusiasmo e o caminho é construído de forma contínua e com trabalho evolutivo.

Para [Grossman and Siegel 2014], quatro questões principais devem ser propostas para que uma evolução analítica seja iniciada e duradoura, sendo elas:

1. As organizações devem olhar os dados como uma função-chave para si e cada unidade de negócio deve perceber a importância da análise de dados e integrá-la às operações.
2. Ter cientistas de dados de alto nível de conhecimento em análise de dados faz com que a compreensão dos problemas a serem resolvidos e as etapas para chegar à solução seja mais clara e precisa.
3. Os cientistas de dados devem conhecer profundamente as unidades de negócio para que seja mais fácil agregar valor a elas.
4. A adoção de uma estrutura de governança analítica ajuda a organização a tomar decisões que priorizam a análise de dados, suas oportunidades e os dados necessários para implementar modelos analíticos.

Ater-se aos elementos da cultura, pessoal, processos e governança (do inglês, *Culture, Staffing, Processes, and Governance - CSPG*), ajuda a estabelecer uma cultura para *big data* e *corpus analíticos*; contratar, treinar e organizar um grupo de funcionários analíticos; desenvolver os processos analíticos necessários; e criar uma estrutura de governança analítica robusta [Grossman and Siegel 2014].

Começando pela cultura, os executivos de nível corporativo devem reconhecer a necessidade de organizar *big data* e *corpus analíticos* como uma função organizacional que recebe ampla responsabilidade e autoridade para ativos de dados e que é análoga a outras funções importantes na organização [Grossman and Siegel 2014].

A estrutura CSPG requer que haja uma massa crítica de equipe analítica (cientistas de dados). A equipe analítica deve ser capaz de obter e gerenciar dados; construir modelos estatísticos, preditivos e de mineração de dados; e implantar esses modelos. O líder de analítica, junto com o gerenciamento corporativo, deve decidir onde localizar a função analítica dentro da organização, podendo ser, essencialmente, a equipe analítica centralizada, descentralizada ou com uma abordagem híbrida [Grossman and Siegel 2014].

O terceiro componente da estrutura CSPG diz respeito aos próprios processos analíticos. O *big data* apresenta muitas oportunidades se esses processos puderem ser criados e gerenciados adequadamente. Os dados podem ser negociados entre as organizações, os produtos podem ser aumentados para produzir dados, os ativos podem ser digitalizados, os dados podem ser combinados dentro e entre setores e assim por diante, quanto mais sofisticados os processos analíticos se tornam, mais oportunidades podem ser exploradas [Grossman and Siegel 2014].

O componente final da estrutura CSPG é a governança analítica. Como *big data* e *corpus analíticos* são novos para muitas organizações, as estruturas de governança analítica não são bem definidas. Os líderes corporativos seniores são responsáveis por definir a estrutura de governança e por monitorá-la e aprimorá-la à medida que a experiência se acumula [Grossman and Siegel 2014].

A governança analítica tem como foco identificar e fornecer recursos para oportunidades analíticas, obter dados que sejam consistentes e oportunos e, por fim, implementar modelos analíticos de forma consistente, a fim de proteger a eficácia das análises [Grossman and Siegel 2014].

A partir da perspectiva CSPG, a jornada de transformação se torna eminentemente humana, a cada novo ciclo de disrupção tecnológica os profissionais e líderes terão oportunidades de se adequar para alcançar novos patamares competitivos [Silva, 2018].

5.1 Análise e discussão dos resultados

Para [Davenport and Harris 2018], alguns elementos fundamentais de um programa analítico de sucesso são compostos por dados, empresa, liderança, metas e analistas que serão subsidiados por dois elementos - tecnologia e técnicas analíticas, necessários para alto desempenho.

Na perspectiva dos autores (*ibidem*, 2018), uma organização pode partir de um estado inerte de deficientes analíticos para um estágio de evolução denominado competidor analítico. Todo o escopo consiste em identificar inicialmente as necessidades de informação, análise e agregar valor aos dados existentes [Teixeira and Valentim 2012].

As organizações estão atualmente em níveis variados de maturidade no que diz respeito à adoção de tecnologias digitais e grau de impacto. As empresas com maior

maturidade são as que enxergam a tecnologia como uma oportunidade e não como um desafio [Poleza e Varvakis 2019].

As alternativas para configurações organizacionais propostas na Seção 2, segue a lógica de que é necessário compreender que uma área de análise deve ser adequada às especificidades organizacionais. Insistir em uma configuração organizacional inadequada não é somente prejudicial para a organização, mas também um fator que pode prejudicar o seu desenvolvimento analítico.

Podemos usar a seguinte pergunta para ajudar a ingressar em um modelo *data-driven*: Como os dados podem ajudar a organização a ser mais competitiva ou sustentável? É importante que as organizações tenham exato conhecimento sobre o que ela quer fazer com os dados, afinal, buscar respostas nos dados para perguntas mal formuladas trará resultados ruins e acarretará perda de investimento e tempo.

Apesar de existirem outras metodologias, a ênfase na metodologia CRISP-DM deve-se ao fato de que é mais abrangente em relação a outras metodologias com o mesmo propósito, cobrindo a fase de conhecimento do negócio e seus objetivos [Sharda, Delen and Turban 2019].

A primeira fase, que busca o entendimento sobre o negócio e seus objetivos, é um diferencial da CRISP-DM quando comparada com as outras metodologias existentes de mineração de dados. O foco no CRISP-DM reforça a percepção de que é necessário aliar os dados ao negócio, pois as coleções de dados são, além de um ativo estratégico, a matéria-prima para a resolução de problemas e busca pelas mais diversas respostas.

Cada estágio evolutivo proposto por [Davenport and Harris 2018] corrobora não somente com a identificação da fase atual e dos requisitos para próximas fases, mas também com a conscientização de que há um longo caminho a ser percorrido para que organizações pouco desenvolvidas analiticamente, como é o caso de organizações *data-informed*, possam alcançar o nível de competidor analítico proposto no estágio 5.

Os investimentos feitos em tecnologia, melhoria de processos, equipes e descoberta de conhecimento podem levar a organização a um novo cenário, onde apenas a intuição dos tomadores de decisão não é mais suficiente e desejável.

Levando em consideração esses aspectos, a organização que estiver mais estruturada em processos, tecnologia, ferramentas, equipe qualificada e uma cultura analítica bem desenvolvida por todo seu negócio, poderá tomar decisões estratégicas com alto grau de assertividade e, assim, se manter competitiva em um mundo de mudanças rápidas.

6. Considerações Finais

As proposições de modelos evolucionários, baseados em estágios de maturidade, contribuem para que uma organização possa mudar de patamar por meio de identificação do seu estado atual de capacidade analítica e fazer uma lista de verificação para atingir o nível superior, como o de competidor analítico de [Davenport and Harris 2018], e ser efetivamente sensível aos dados processados.

Assim como os estágios de evolução que foram apresentados, os modelos de configuração organizacional contribuem para que a organização saiba posicionar a área de ciência de dados a ser implantada, levando em consideração as características peculiares da organização.

A metodologia de análise de dados, CRISP-DM, auxilia no entendimento de como os dados devem responder às necessidades do negócio, contribui para caracterizar uma organização *data-driven*, e os conceitos abordados servem como contribuição para o processo de avanço de *data-informed* para *data-driven*.

Com isso, o objetivo do trabalho foi atingido na medida em que explorou aspectos conceituais e categóricos no âmbito organizacional, buscando identificar possíveis relações de modelos de configuração organizacional e abordagens metodológicas de análise exploratória sobre dados com fatores habilitadores para uma jornada *data-driven*.

Baseado na obra de [Davenport e Harris 2018], assume-se a hipótese de que, independente do estágio em que a organização se encontra, é importante focar em alguns itens para se tornar analiticamente competitiva, a saber: coleções de dados, ao processá-las pode-se revelar informações que subsidiam as decisões e a produção de conhecimentos valiosos; empreendimento, ao gerenciar e coordenar recursos a nível empresarial; liderança, ao promover um time e uma cultura de liderança analítica; alvos, ao focar investimentos analíticos nas áreas promissoras e de maior retorno; analistas, ao desenvolver e gerenciar talentos analíticos.

Para conseguir evoluir para um modelo *data-driven*, a organização precisa mudar as suas práticas e processos. Diferentes cenários de implementação podem ser vislumbrados como, por exemplo, ambientes organizacionais desprovidos de modelos sensíveis a dados, cuja metodologia de análise de dados abordada e outras estratégias de implementação devam ser mapeadas, a depender do contexto e da necessidade de negócio.

Portanto, as organizações precisam ter consciência da importância dos negócios amparados pelas tecnologias da informação e de dados, quanto valor podem gerar seguindo um modelo centrado em preceitos analíticos e quais as respostas para os problemas de determinados domínios de conhecimento. A organização poderá explorar áreas de domínios distintos quando conseguir compreender a importância de ser orientada a dados ao invés de utilizar apenas para um suporte de apoio e gestão.

Referências

- Anderson, Carl. (2015). "Creating a data driven organization", Sebastopol: O'Reilly Media.
- Camargo, R. F. (2021). "A importância dos Tipos de Estrutura Organizacional para um negócio de sucesso", Treasy, 2016. Disponível em:

- <https://www.treasy.com.br/blog/tipos-de-estrutura-organizacional/>. Acesso em 11 maio 2021.
- Davenport, T. H. and Harris, J. G. (2018). “Competição analítica: vencendo através da nova ciência”, Rio de Janeiro: Alta Books.
- Griffin, Jane and Davenport, Tom (2014). “Organizing Analytics - From the inside out”, Deloitte Analytics. Disponível em: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/deloitte-analytics/us-da-organizing-analytics-inside-out.pdf>. Acesso em 13 maio 2021.
- Grossman, Robert L. and Siegel, Kevin P. (2014). “Big Data and Organization Design”, Journal of Organization Design. Vol. 3, n. 1, p. 20-24.
- Kampakis, Stylianos. (2021). “What are the differences between data-driven, data-informed and data-centric”, The Data Scientist, [S.I.], [2020?]. Disponível em: [https://thedata scientist.com/data-driven-data-informed-data-centric/#:~:text=Data%20informed%20companies%20are%20collecting,the%20core%20\(e.g.%20Google\)](https://thedata scientist.com/data-driven-data-informed-data-centric/#:~:text=Data%20informed%20companies%20are%20collecting,the%20core%20(e.g.%20Google)). Acesso em 15 maio 2021.
- Kelleher, John D. and Tierney, Brendan. (2018). “Data Science”, 1. Ed. Cambridge: The MIT Press.
- Matias, Alberto Borges (2021). “Tipos de estrutura organizacional: tire suas principais dúvidas”, Inepad Consulting, 2018. Disponível em: <https://blog.inepadconsulting.com.br/tipos-de-estrutura-organizacional/>>. Acesso em 11 maio 2021.
- Miranda, Gloria Macias-Lizaso. (2018). “Special Edition on Advanced Analytics in Banking”, McKinsey on Payments. vol. 11, n. 28, p. 1-8, ago.
- Poleza, M. and Varvakis, G. (2019). “Processo de transformação digital nas organizações: razões e resultados”.
- Provost, Foster and Fawcett, Tom. (2016). “Data Science para Negócios: o que você precisa saber sobre mineração de dados e pensamento analítico de dados”, 1ed. Rio de Janeiro: Alta Books.
- Rautenberg, Sandro and Carmo, Paulo Ricardo Viviurka. (2019). “Big Data e Ciência de Dados: complementariedade conceitual no processo de tomada de decisão”, Brazilian Journal of Information Studies: Research Trends. 13:1 p.56-p.67. ISSN 1981-1640.
- Sharda, Ramesh; Delen, Dursun and Turban, Efraim. (2019). “Business Intelligence e Análise de Dados para gestão do negócio”, Tradução de Ronald Saraiva de Menezes. 4. Ed. Porto Alegre: bookman.
- Silva, Nelson. (2018). “Transformação Digital, a 4ª revolução industrial”, Caderno opinião - FGV Energia.

- Stewart, Shayna. (2019). “Are You Data-driven, Data-informed or Data-inspired?”, Amplitude, 7 de jun. Disponível em: <https://amplitude.com/blog/data-driven-data-informed-data-inspired>. Acesso em 15 maio 2021.
- Teixeira, Fernando. (2021). “DDOM: os 6 pilares de uma empresa data-driven”, Mit Sloan Management Review Brasil. Disponível em: https://www.mitsloanreview.com.br/post/os-6-pilares-de-uma-empresa-data-driven-o-u-ddom?gclid=CjwKCAiAuoqABhAsEiwAdSkVVGFWVINLmI0hE8kPApLYFak1KFfatMWJn_EMVlaKewWajN5XNwc_dxoC4-AQAvD_BwE. Acesso em 05 jun. 2021.
- Teixeira, Thiciane M. C. and Valentim, Marta L. P. (2012). “Estratégias para Disseminação do Conhecimento Organizacional: o papel da arquitetura da informação; Estratégias para la Diseminación del Conocimiento Organizacional: el papel de la arquitectura de la información”, Informação & Informação, v. 17, n. 3, p. 165–180, 11 jan.