



Google Cloud Cast

## Episódio #2: Tendências para líderes de Data e Analytics



### Daniel Leite, Executivo de Vendas do Google Cloud

Descrição da imagem: foto com o rosto de Daniel Leite. Daniel tem pele clara, cabelos curtos e na foto aparece usando uma camisa social e sorrindo.



### Marcelo Gomes, Executivo de Vendas do Google Cloud

Descrição da imagem: foto com rosto de Marcelo Gomes. Marcelo tem pele clara, barba curta, usa óculos e na foto aparece de camiseta preta e sorrindo.



### Michel Avila, CDO da Neoway

Descrição da imagem: Foto com rosto de Michel Avila. Michel tem pele clara, cabelo curto e escuro, barba curta e na foto aparece sorrindo e de camisa de botões.



**Daniel Leite:** Olá, sejam todos bem-vindos a mais um episódio do Google Cloud Cast, o nosso podcast preferido aqui do time de Google Cloud. Eu sou Daniel Leite, executivo de vendas de Google Cloud. Estou aqui com meu parceiro de podcast, Marcelinho. Marcelinho, tudo bem contigo?



**Marcelo Gomes:** Tudo joia, Dani? Eu sou Marcelo Gomes. Eu também sou executivo de vendas aqui no time de Google Cloud. E, caso vocês queiram entrar em contato com a gente, lembrem-se sempre de que nós temos o e-mail [googleclouddcast@google.com](mailto:googleclouddcast@google.com). E hoje nós temos um convidado especial aqui, certo, Dani?



**Daniel Leite:** Exatamente. Um convidado especial para um episódio mais do que especial. Para os nossos queridos ouvintes, acho que vale a pena fazer uma introdução antes de trazer o nosso convidado para conversar conosco. Vamos lá!

Começando com um dado, uma pesquisa realizada pela Harvard Business Review em 2020 mostrou que os líderes das organizações são os que mais valorizam as ferramentas de gestão e também são os que mais acreditam que o acesso a esses dados deve ser democrático dentro da organização. Isso para garantir, claro, bons resultados e o sucesso do negócio.

Outra fonte, outra pesquisa realizada pela Gartner, uma grande empresa de consultoria de TI, mapeou as tendências do setor para 2022 (este ano em que estamos) e revelou que, dentre todas as prioridades dos líderes, *data* e *analytics*, ou seja, dados e analytics aparecem em segundo lugar, perdendo apenas para e-business, que nada mais é do que a digitalização dos negócios das empresas.



**Marcelo Gomes:** Ou seja, os tomadores de decisões estão falando sobre dados e análise de dados, e nós não poderíamos deixar de trazer esse tema para o nosso podcast, certo? Então, neste nosso episódio, nós vamos falar sobre as principais tendências de *data* e *analytics* para 2022, as oportunidades e os desafios enfrentados pelos líderes dessas áreas, e como o Google Cloud pode fornecer recursos e soluções para ajudar a facilitar esse processo de transformação digital das empresas



**Daniel Leite:** Exatamente. Já que fizemos a introdução de um tema amplo e que dá para discutir bastante, vamos receber o nosso grande convidado de hoje, o Michel Ávila, que é CDO da Neoway, e conversar um pouco mais conosco sobre não só as tendências, mas as diversificações que a gente pode ter dentro das empresas para o futuro. Michel, seja bem-vindo!



**Michel Avila:** Oi, Dani! Oi, Marcelo! Antes de mais nada, parabéns pelo programa. Está sensacional mesmo! E obrigado pela oportunidade de bater esse papo aqui com vocês hoje.



**Marcelo Gomes:** A gente que agradece, Michel.

E eu vou começar já com a primeira pergunta. É muito notável que, nos dias de hoje, [os dados] estão inseridos em todas as áreas de uma empresa, todo o cenário de dados. Eles são necessários, com certeza, para a tomada de qualquer decisão. E eles ocupam uma posição muito estratégica dentro das empresas. E para a gente começar esse nosso papo, acho que a gente pode começar falando um pouco sobre esse tema de como os dados se relacionam com as áreas dentro de uma empresa e por que os líderes enxergam tanto valor neles a ponto de tornar toda essa parte de análise de dados uma prioridade.



**Michel Avila:** Legal. Olha, tem uma frase que eu acho perfeita sempre para abrir esse tipo de papo, que é: “Em Deus nós confiamos. Os demais devem trazer dados”. Só que eu estava dando uma pesquisada nessa frase e ela já era um clichê na comunidade científica há séculos, porque, obviamente, ela faz parte do *mindset* de todo cientista. Então a *buzzword* “data-driven” não traz nenhum conceito novo em si. Ela é só um apelo para que as empresas adotem esse mesmo *mindset* em larga escala. Até o dia em que colocar os dados à frente das opiniões seja uma parada bem mais cultural. Isso não só para as áreas de negócio, mas também para todas as áreas da empresa. Não só na hora de fazer um grande planejamento estratégico para os próximos anos, mas principalmente no dia a dia. Só que esse tema — já respondendo a pergunta — só ganhou força e prioridade para as lideranças das empresas porque tem muita grana na mesa. Só aqui na Neoway, nesses 12 anos de trabalho, eu vi centenas de cases de sucesso com retornos vertiginosos para os mais variados tipos e tamanhos de negócio. É uma posição privilegiada, claro, porque eu acompanho não só a evolução da maturidade da minha empresa, mas das principais empresas do país inteiro, à medida que elas vão entendendo seus dados, vão dominando esse cenário e trazendo demandas cada vez mais complexas.

E, falando nisso, eu trouxe um case. A gente, em algum momento, vai falar mais, vai tentar trazer alguns cases aqui, mas eu trouxe um case que mais parece um “causo” de tão bem que ele ilustra o choque de realidade que o uso intensivo de dados pode causar numa empresa. Então vou contar para vocês rapidinho. Lá em 2010, o dono de uma fábrica de revestimentos cerâmicos fez uma pergunta muito simples para a gente: quanto eu devo produzir esse ano? Se eu produzir a mais, eu vou ter que torrar o excedente no mercado para pelo menos cobrir o custo da produção. Se eu produzir a menos, eu deixo de vender tudo o que eu poderia e ainda abro espaço para a concorrência. Então, eu não posso errar na hora de programar a fábrica. Na época, a principal informação utilizada para estimar essa demanda da fábrica era o número de obras mapeadas pelo próprio time de vendas, que fazia isso manualmente. Então o dono da empresa perguntou para o gerente de vendas: “Me fala quantas obras a gente tem em um raio de 8 km da nossa sede?”. Ou seja, no quintal da casa deles. O gerente saiu num “barata voa”, catando umas planilhas, umas cadernetas, papel daqui,

papel dali, e chegou a um total de 18 obras. Aí o dono virou para a gente e fez a mesma pergunta. O nosso sistema, a gente foi lá, pesquisou no nosso sistema e ele informou 120 obras contra 18. Óbvio, acabou a reunião, ficou um climão e o gerente diz que o nosso sistema estava, obviamente, todo “bugado”, que aquele número não tinha nenhum sentido. E o que a gente fez? No dia seguinte, a gente pegou um *motoboy* para ir de obra em obra e tirar uma foto de cada uma delas. Conclusão: das 120 obras, 112 estavam realmente ativas e eram dentro do perfil da fábrica. As demais só não iriam precisar mais de cerâmica. Já tinham passado da fase da cerâmica. Então, agora pensa: se a empresa não conhecia nem o mercado ao redor da sede, no quintal deles, imagina o cenário das localidades que são mais remotas, em outros estados.

Então aqui a gente tem um exemplo muito claro, quase uma caricatura, do que acontece com uma empresa quando ela não se apoia em dados de verdade para tomar decisões, especialmente as decisões mais críticas — fabricar ou programar uma fábrica para o ano inteiro. E também um exemplo de que cedo ou tarde vai ser necessário buscar dados também fora da empresa. Nem todos os dados de que a gente precisa, às vezes, para tomar essas decisões estão internamente. Então, conhecendo todas as obras ativas, dava para visitar, negociar volume, programar a fábrica com segurança e ainda trazer alguma previsibilidade de receita, além de aumentar o mercado de atuação. Então, óbvio, foi um retorno gigantesco para a fábrica e ficou ali um aprendizado enorme para a gente também, porque nem a gente tinha ideia do impacto que uma busca por dados e uma estruturação de uma decisão dessas usando dados teria um impacto tão grande.



**Daniel Leite:** Aí você trouxe, Michel, um case já mais profundo, já bem aterrizado do nosso mundo real, que é: às vezes o dado está disponível, a gente só não consegue organizá-los e consumi-los de maneira muito específica para fazer, por exemplo, alavancar um negócio como era o caso dessa fábrica que você comentou.

Já puxando esse mesmo gancho dos desafios por que as empresas passam hoje no dia a dia com tecnologia, quando a gente olha para esses desafios, claro, é um mundo muito discutido dentro das empresas de tecnologia que vão lidar com isso no dia a dia, tem empresas ainda de outros setores que estão em uma fase inicial para saber como trabalhar com esses dados que, com certeza, vão demandar muitas análises para ferramentas e plataformas como a de Google Cloud, que você conhece muito bem. Minha pergunta vai nesse sentido: por que é tão importante as empresas investirem nessas ferramentas e quais são os benefícios que elas vão ter no final da jornada para essas decisões estratégicas, para esse alinhamento mais correto para o negócio delas?



**Michel Avila:** Legal. Olha, eu vejo isso como um desafio, no final das contas, de escala. Essa busca por eficiência operacional, por informações de qualidade, para tomar as melhores decisões é super antiga. Até ontem, algumas empresas com mais, outras com menos sofisticação, elas davam um jeito de coletar algum dado daqui e dali, organizar, juntar, analisar e extrair dali alguma inteligência. Só que o que acontece? Mais cedo ou mais tarde essas técnicas deixaram de funcionar porque o volume de dados cresceu demais ou porque a

frequência deles ficou muito alta ou — o que é um cenário que a gente tem mais frequentemente hoje — essas informações acabaram tomando formatos muito variados, então a gente tem informações que estão em planilhas, outras em documentos, tem imagem, tem áudio, tem vídeo, agora tem sinais de IoT, sinais de máquina.

Então, respondendo diretamente à pergunta, muito antes, eu acho, de ser um problema de alfabetização de dados (o *data literacy*), as empresas têm dificuldade de extrair inteligência dos seus dados porque elas não têm as ferramentas adequadas, exatamente como você colocou. Então, esse movimento de transformação digital é isso, no final das contas. Ele marca esse momento crítico em que uma empresa esbarra nesse muro, percebe que ela precisa evoluir e decide iniciar uma jornada de digitalização. Então, no final do dia, o objetivo é redescobrir a capacidade, recuperar a capacidade de produzir inteligência dentro desse novo cenário. E aí, diante dessa situação, é natural escolher uma cloud que já venha recheada de soluções prontas para sair usando, ofereça uma estratégia de custo que varia de acordo com o consumo, o que é super relevante. Às vezes, a empresa está tímida, não sabe se vai, se não vai. Aí vai fazer uma cotação e, de repente, sai um negócio supercaro. Já fica proibitivo. Então essa estrutura de custo que varia de acordo com o consumo é superconvidativa e zero preocupação com infraestrutura, datacenter, segurança física, refrigeração. Enfim, só quem já trabalhou em empresas um pouco mais antigas (estou entregando a idade) sabe o que é toda a manutenção, toda a operação, tudo o que ocupa de tempo e de grana cuidar de um datacenter, cuidar de máquinas físicas.

Então, uma vez dentro de uma *cloud*, a facilidade de subir dados, de montar os primeiros relatórios, já superempolga. E aí, como o mato está muito alto nesse início, mesmo os relatórios mais simples que você consegue montar já têm um impacto enorme, porque eles conseguem ser facilmente democratizados e acessados por várias áreas diferentes. Então o início, digamos assim, dessa *cloud* não é fácil, mas também empolga bastante. E aí, claro, é um caminho sem volta. Nunca vi ninguém que foi para a nuvem e depois se arrependeu e montou o seu próprio datacenter. Tem algumas empresas que fazem isso, mas são características, são negócios muito, muito específicos. Então esse é um caminho sem volta. E aí, à medida que a empresa vai utilizando esses relatórios, vai adotando esses relatórios no dia a dia, os usuários vão ficando mais exigentes. E aí eles começam a questionar um monte de coisa. A gente começa a evoluir, começa a perguntar de onde está vindo aquela informação, quando foi atualizada pela última vez, como um determinado índice está sendo calculado, com quem eu falo para entender melhor uma determinada informação. E se essas perguntas ficam sem resposta ou se demora demais para conseguir, ou pior, se elas mudam de acordo com quem responde, o programa começa a perder credibilidade e ir por água abaixo.

E é justamente aí que entra a governança de dados. Ela vem para colocar ordem nessa casa, é um facho de luz no meio desse breu de dados. E, como não poderia ser diferente, a nuvem do Google traz a peça fundamental desse ecossistema, a pedra angular, que é o catálogo de dados. Sem isso, não tem governança, não tem transparência, não tem confiança. É caos total. Se fôssemos fazer uma analogia, seria como abrir um supermercado sem saber o que tem na loja nem no estoque. Então é o catálogo que lista todos os itens de dados disponíveis, que informa a origem de cada um, que dá informações técnicas e de negócio desses dados, informa a pessoa, o time ou setor

responsável, se precisar entender melhor aquilo. Claro que isso não vem pronto, mas quando você combina esse catálogo com uma BigQuery, quando você pega um gestor de acessos e identidade — o IAM — vira um ecossistema inteiro superpoderoso para auxiliar nessa jornada, e não só de fazer a gestão desse volume gigantesco de dados, mas também garantir que todos os relatórios, toda essa adoção da nuvem não vá por água abaixo, que você vai conseguir dar todas as respostas, dar essa segurança que a galera precisa para continuar usando e se desenvolver ainda mais. Lá para a frente, quando a gente falar sobre proteção e privacidade de dados, a gente de alguma forma vai conectar de novo com esse catálogo que eu acabei de falar.



**Marcelo Gomes:** Poxa, Michel, que resposta super completa! Foi legal que você passou por vários pontos. Inclusive você deu um contexto do que vocês auxiliam dentro do contexto de organização e coleta de dados, interpretação dos dados, até o ponto de a gente poder tratar esses dados. E o catálogo de que você falou é um dos pontos-chave. Mas eu queria voltar a um ponto que você falou, que você citou no meio da sua resposta. Pensando ainda no contexto de interpretação de dados, acho que está no momento de citar a questão da tendência que você falou, que grandes players do mercado já estão até utilizando e implantando, que é o *data literacy*, que no cenário a gente chama de um termo traduzido que seria a alfabetização dos dados. Você consegue falar um pouquinho mais sobre isso para a gente? Por que seria uma tendência esse contexto?



**Michel Avila:** Claro, com certeza. Esse tema sempre me lembra o momento em que os computadores realmente invadiram o mercado e passaram a ser empregados em larga escala. Eu lembro daquele momento em que a gente precisou passar mesmo por uma fase de alfabetização em informática. Um monte de gente fez cursinho de informática porque não adiantava ter um computador à disposição se ninguém sabia usar aquela geringonça. Curiosamente, essa fase tem uma conexão direta com esse tema. Foi justamente depois que a gente dominou essas máquinas que a produção de dados disparou de forma descontrolada. E hoje a gente tem um monte de dados espalhados dentro e fora das empresas e pouca gente sabe o que fazer com eles, como a galera com um monte de computadores que não sabia nem ligar naquela época.

E a importância dessa fase é a alfabetização em dados. É essa alfabetização que vai nos dar o conhecimento básico necessário para navegar nesse novo mundo, assim como foi lá atrás aprender a navegar pela internet. Essa alfabetização — para a gente não ficar só em uma parte mais superficial — a grosso modo, ela consiste em estudar o ciclo de vida dos dados e todas as atividades que acontecem em cada etapa dessa. É você entender como nasce, vive e morre um ativo de dado.

A primeira etapa, obviamente, é a origem do dado. Também é a mais importante. A pergunta número um em qualquer empresa, especialmente quando há algum problema, é: de onde veio essa informação. É fundamental a gente saber se aqueles dados nascem em um sistema, se eles foram informados por uma pessoa, se eles foram coletados diretamente de alguma fonte. E aí uma pessoa letrada em dados entende as diferentes origens dos

dados e sabe determinar o nível de confiança e o valor de cada uma delas para o negócio. Lembrando sempre que tem um pano de fundo para essa jornada dos dados que é o negócio. A gente está tentando fazer isso para atender a algum problema, para criar alguma solução para o mercado.

A segunda etapa vai ser o entendimento desses dados. O que eles representam no mundo real, os conceitos principais relacionados a eles, assim como o formato, que determina o nível de dificuldade na hora de manipular. Extrair dados de um arquivo em CSV é muito mais fácil do que de um texto livre e de um PDF, em que a complexidade é bem maior. Então essa análise que a gente faz na segunda etapa dá uma ideia inicial dos dados que podem ter mais ou menos valor e que podem ser analisados primeiro. E uma pessoa letrada em dados conhece os diversos formatos que um dado pode assumir e entende os desafios de manipular cada um deles. Então a segunda etapa é quando você olha e separa o joio do trigo, e entende um pouco daqueles dados que estão disponíveis para fazer alguma coisa.

Na terceira etapa, a gente tem a ingestão disso, a higienização, a padronização, o enriquecimento e a organização desses dados. Então aqui é onde rola a ETL clássica, o *data jujitsu*. A grande e esmagadora parte do tempo que a gente investe é nessa parte, porque aqui é onde a gente tem a maior dificuldade mesmo, o maior desafio. E uma pessoa que é letrada em dados conhece as diferentes técnicas e as ferramentas que podem ser aplicadas aqui. E, olha, são inúmeras! Essa [etapa] três dá um minicursinho fácil, fácil.

Na quarta etapa, os dados já estão prontos para ser analisados. Então aqui vale tudo: pode combinar, separar, filtrar, agrupar, gerar novos dados. O objetivo é compreender algum fenômeno, investigar alguma concorrência, testar uma hipótese e criar versões que possam ser mostradas em painéis e relatórios. Uma pessoa que é letrada em dados conhece as ferramentas de apresentação visual mais usuais e sabe que cada tipo de análise vai exigir uma organização diferente dos dados.

A quinta etapa é quando a maturidade no manejo dos dados já está elevada o suficiente e já é consistente. Então é aqui que começa a análise dos dados históricos em busca de tendências. O *analytics* de que a gente fala tanto vai começar aqui na etapa 5. Então aqui a gente vai pegar correlações, previsões, até sugestões de ação — os modelos preditivos, os modelos prescritivos. Uma pessoa que é letrada em dados sabe que existem abordagens que vão combinar estatística, computação, regras de negócio, para produzir indicadores de probabilidade e vão gerar *score*, previsões e até as recomendações. É aqui que aquelas montanhas — que a gente pega uma montanha de dados — e, no finalzinho, ela gera um “sim” ou “não”. Onde a gente transforma o big data em *smart data*.

Teria acabado aqui mas, graças à nossa querida LGPD, a quinta etapa não é mais a última. Agora a gente tem a sexta etapa, que é um dos temas mais quentes no mundo jurídico e de dados nos últimos dois anos: a garantia da proteção e privacidade dos dados é quem mais clama por essa alfabetização. Hoje, quem mais puxa isso é a própria LGPD. Então, uma pessoa que é letrada em dados entende os aspectos jurídicos básicos relacionados ao tratamento de dados de pessoas físicas e todas as responsabilidades envolvidas. Ela sabe que os dados precisam ter uma origem legítima, que não podem ser utilizados para qualquer finalidade, e também não podem ser armazenados para sempre.

Então, claro que conhecer esse ciclo de vida e saber o que acontece em cada etapa é diferente de saber executar cada uma delas. O desafio é gigante. Então essa alfabetização em dados é superficial e generalista mesmo, justamente para que ela possa ser abrangente. Não sei se isso responde à sua pergunta..



**Daniel Leite:** Com certeza. Pensando nas cinco ou, no caso, seis etapas, que o Michel cita, tem tanta coisa que a gente tem que realmente pensar e estruturar para chegar do outro lado, né? Mas foi muito bom o exemplo. A gente vai usar o *roadmap*, o caminho que tem que ser seguido pelo uso do melhor dado ou como consumir conscientemente. E eu quero te perguntar bem nesse ponto, partindo da alfabetização dos dados. É claro que é esperada a tomada de decisões baseadas em dados e que ela aconteça de uma forma muito mais rápida, muito mais assertiva. Eu acredito que daí vem uma segunda grande tendência, que é a criação dos componentes, ferramentas, softwares, tudo o que interage com esses dados, para que eles possam agir de forma a agilizar a interpretação dos dados, transformá-los em ações, como você comentou e, claro, se não for possível fazer isso dentro da própria empresa, terceirizar esse serviço, usando, por exemplo, a plataforma de Google Cloud. Acho que esse caminho é muito claro, não é, Michel?



**Michel Avila:** Sem dúvida. É um processo evolutivo, não tem como ser diferente. Quando a gente começou a usar as plataformas em nuvem para processar grandes volumes de dados, a gente tinha dois trabalhos enormes.

Primeiro, mandava tudo para a nuvem. Jogava os dados todos para a nuvem e aí organizava lá os recursos computacionais, rodava os MapReduce, os Spark, que era o jeito que a gente conseguia fazer alguns grandes processamentos realmente acontecerem. Mas, quando terminava, a gente tinha que baixar tudo. Terminou de processar, tinha que baixar tudo de novo de volta da nuvem para algum outro lugar. Só que, pensando na experiência do dado em si, isso gera uma ruptura no fluxo de vida da informação. Esse “sobre para a nuvem, desce da nuvem” é um negócio que gera uma quebra de contexto. O ideal era que os dados já estivessem lá e que também não precisassem sair de lá. Então, pensando em mercado, tinha uma oportunidade enorme ali para ser explorada.

E aquilo que começou como uma solução só para orquestrar recursos computacionais — que, no caso, até estavam ociosos — e permitir o processamento de grandes volumes de dados, inevitavelmente teve que evoluir muito rápido para acomodar essas outras fases do ciclo de vida. Com isso, deixa de ser necessário baixar os dados da nuvem para terminar o trabalho em algum outro lugar. Você pode conectar esses dados a uma ferramenta, por exemplo, como o Google Data Studio, e já passar para a próxima etapa. Seria uma etapa de fazer uma análise, uma etapa de entender os dados, de ver os resultados. Essas ferramentas são pensadas para usuários não técnicos, para que seja mesmo rápido, fácil e intuitivo criar um dashboard, gerar um relatório e desenhar um estudo. Isso permite entregar aos decisores da empresa informações que são importantes e de forma muito rápida. E você faz isso e essa experiência gera um feedback superpositivo e aí faz com que mais e mais

relatórios e dashboards sejam encomendados. E, assim, a empresa acaba embarcando nesse ciclo virtuoso das decisões baseadas em dados, do *data-driven*, em que você toma uma decisão, vê que dá um resultado positivo, se sente mais confiante de tomar as próximas, e aí você busca mais embasamento em dados. Esse é o caminho para a maior parte das empresas, porque construir tudo isso dentro de casa é um investimento que nunca vai se pagar, além de ser um desvio enorme de foco. Então, olhando para a nuvem hoje, tem um monte, um monte, um monte de pequenas ferramentas e soluções que já estão ali dentro que permitem fazer esse trabalho com os dados de fim a fim, lá do iníciozinho até o fim, até você entregar uma decisão, entregar uma inteligência, uma avaliação, sem ter que causar uma ruptura nos dados.



**Marcelo Gomes:** Muito legal, Michel. Você falou bastante de todo o ciclo do dado, de a gente gerar no final um dashboard. Nada melhor do que a gente conseguir ver no final um resultado e, se a gente puder ver de uma forma mais simples de interpretação, eu acho que facilita a vida de todo mundo, principalmente para quem for tomar uma decisão rápida.



**Daniel Leite:** Vou te cortar só para fazer um comentário. Você viu que a minha pergunta foi bem orientada e, no final, a gente potencializou as nossas soluções de Google Cloud! [risos]



**Marcelo Gomes:** Com certeza, com certeza. E o legal, se a gente pensar, é que o Data Studio ainda é uma solução que não tem custo. Então ele consegue conectar em outras soluções e gerar insights de uma forma até mais otimizada (vendendo o peixe aqui, aproveitando).

Eu queria aproveitar até para citar uma outra tendência, apesar de não ser necessariamente nova, que é a reengenharia. Não sei se todo mundo que nos ouve, nossos queridos ouvintes, estão acostumados com esse termo, mas é basicamente a reestruturação dos processos de trabalho por força das novas condições de mercado. Ou seja, o objetivo da reengenharia é justamente a gente conseguir encontrar uma maneira melhor de realizar os processos buscando melhores resultados. Como o próprio Michel comentou, a LGPD veio para trazer uma necessidade de acompanhamento de todos esses dados. Isso vai de encontro com o que a gente falou até agora dentro do nosso podcast, que é a gente conseguir utilizar os dados para melhorar a tomada de decisões e ser mais rápido, ser mais ágil para fazer isso e ter, com certeza, assertividade. O Michel trouxe um exemplo logo no início que ilustra tudo isso: que às vezes a gente ter os dados não quer dizer que a gente tenha os dados da melhor forma ou todos os dados de que a gente precisa.

A pergunta em si para você, Michel: como os líderes podem colocar a reengenharia em prática quando a gente fala no cenário de análise de dados?



**Michel Avila:** Boa! Essa tem um pouco mais a ver com o *data-driven*. Essa aqui é de 1990, a reengenharia. Ela provoca mesmo uma reflexão mais profunda sobre o jeito de se fazer as coisas dentro de uma empresa. E quando se fala em mudança, já nasce uma resistência natural por parte das pessoas, seja por medo de ter que reaprender uma atividade que já tinham dominado ou serem realocadas para outra função ou, ainda pior, se tornarem obsoletas dentro daquela empresa. Esse é um fato que a gente não pode ignorar quando traz um tema desses para a mesa. Ao mesmo tempo, os casos de uso — isso eu gosto de falar bastante — os casos de uso de inteligência artificial mais contundentes que existem sempre incluem o elemento humano no processo. Então o segredo é encontrar essa combinação perfeita.

Para começar, os processos que são diretamente ligados aos produtos e serviços fornecidos pela empresa devem ser avaliados primeiro, obviamente, porque eles vão ter muito mais chance de causar um impacto relevante do que os demais. Uma outra dica são aqueles processos que não escalam. Ou seja, se você quiser o dobro de resultado, eu preciso do dobro de pessoas. Geralmente, esse problema de escala a gente consegue resolver fazendo alguma reengenharia nesse processo e encontrando alguma coisinha ali dentro que dá para a gente automatizar ou pelo menos aumentar com alguma inteligência. E, por último, mas não menos importante, aqueles processos que envolvem análise de dados e tomada de decisão. Só nessas três categorias já vai ter um monte de processo pré-selecionado para analisar primeiro. Se a gente entra numa pegada de reengenharia, a primeira pergunta é: em quais processos eu vou fazer a reengenharia primeiro? Então, dessas três dicas que eu dou aqui, já dá para ter alguma ideia de uma primeira seleção.

Dentro desses, uma vez dentro do processo em si — você vai olhar os passos, como é que funciona, quem faz o quê, quais são as atividades — dá para atacar aquelas atividades que são muito repetitivas e que não requerem muito raciocínio, por exemplo, porque a repetição de atividades, atividades muito repetitivas e que não precisa de muito raciocínio, são perfeitas para a gente fazer isso com máquinas, perfeitas para a gente tentar descobrir alguma forma de automatizar isso. E a outra boa pedida é avaliar se tem alguma etapa de decisão que poderia ser melhorada ou até mesmo delegada para uma inteligência artificial. Dessa forma, a gente deixa para a máquina as partes que são mais entediadas dos processos e libera as pessoas para pensar em formas mais eficientes de chegar ao mesmo resultado. Para mim, esse processo de rever o *status quo* deveria ser uma prática contínua da empresa, como um corpo que renova as suas células constantemente. E, como não dá para falar de reengenharia sem fazer essa provocação, aqui vai, não tem como: não raro a gente busca um cavalo mais rápido quando a solução é um carro mesmo. Então, a reengenharia, para mim, sempre tem essa pitadinha de inovação disruptiva que sempre vai bem, para a gente não cair na tentação de digitalizar um processo que é feito de forma manual, sabe? A gente tem hoje — especialmente com a tecnologia que a gente tem, com o ambiente da nuvem, as *clouds* — a gente tem oportunidades reais de transformar a forma como a gente faz as coisas. Então a reengenharia traz esse convite, mas ela tem essa... Se você não prestar atenção, você começa simplesmente a fazer digital aquilo que você fazia no manual antes. E a gente tem que sempre ter uma pitadinha de inovação disruptiva para ver se a gente não pode fazer de um jeito totalmente diferente aquilo ali e obter o mesmo resultado mais rápido, com mais escala e mais barato também.



**Daniel Leite:** Você comentou, Michel — além, claro, de tudo o que você falou — você colocou algumas palavrinhas-chave. Vou até encaixar a minha próxima pergunta voltando a um dos temas de que você falou lá atrás, que foi a sua quinta estendida ou a sexta parte específica da explicação sobre dados, sobre a jornada de dados. A gente sabe hoje que uma tendência — que já nem é mais tendência, é uma realidade de fato — quando a gente olha para os dados, para a análise, é a cyber segurança. Então você citou LGPD, que é a LGPD aqui no Brasil, GDPR na Europa e a PIPL na China, que são todas as normativas que olham para esse tipo de segurança do dado, e a gente, claro, pode afirmar que é um movimento de proteção que é uma prioridade e não se trata mais de simplesmente gerenciar os dados, mas sim como as empresas vão proteger e se adequar à lei para não terem nenhuma sanção, nenhum problema específico com os dados dos seus usuários e, no final, causar um problema maior para si mesmas. Aqui, um ponto específico: vale lembrar que o Brasil é um dos países com maior número de vítimas de ataques cibernéticos comparado a todos os outros. Quem não se lembra, por exemplo, no levantamento do podcast, um exemplo (mas gritante!) foi o megavazamento que aconteceu de 220 milhões de pessoas no começo do ano passado aqui no Brasil. Então é por esse tipo de coisa que a gente realmente se preocupa. Cyber segurança é um ponto que olha especificamente para isso.

Acho que agora eu queria dar uma mergulhada um pouco mais. Queria que você ajudasse a gente, os nossos ouvintes, para explicar um pouquinho melhor quais são esses maiores desafios das empresas em atender a essas novas regulamentações? Quais são as soluções de proteção, de segurança, que plataformas como a de Google Cloud podem oferecer para a gestão de dados nas empresas? Claro que isso é um guarda-chuva bem grande, mas eu queria que você começasse essa resposta e depois entrasse no ponto das facilidades que os clientes podem esperar também nesse uso das soluções em nuvem.



**Michel Avila:** Show! 220 milhões de pessoas naquele vazamento é a base inteira de pessoas físicas do Brasil. Então vazou “só” a base inteira. Mas, enfim, tem uma frase que a gente ouviu durante algum tempo: “dados são o novo petróleo”. E aí depois virou, e teve “n” variações: “dados são pessoas”. Eu ouvi uma vez uma que [diz]: “dados são o novo bacon”. Atualmente, a minha preferida é: “dados são o novo elemento radioativo”, porque é um paralelo perfeito para dar a dimensão do que é tratar dados pessoais, especialmente quando a gente pensa em vazamento de dados.

Lembra ontem, quando vi uma notícia de vazamento de dados, era só para dizer que a empresa não investiu em segurança, meter o pau na empresa e vida que segue. Hoje, depois da LGPD, cada vazamento de dados é uma mini Chernobyl. Primeiro que tem que notificar até o Papa sobre o incidente e depois, enquanto para o time de segurança da informação aquilo contabiliza um incidente, agora, para a ANPD, é um incidente, uma multa, para cada pessoa vazada! Ou seja, são multas potencialmente milionárias. Então, olhando somente para segurança, vamos separar esse aspecto de proteção e privacidade de dados em duas grandes linhas. Uma linha é a questão da segurança: preciso garantir que esses dados não vão vazarem, que não vai ter um acesso ilegítimo a essas informações.

A LGPD amplificou muito o peso e os impactos desses vazamentos. Então está aí uma parte do desafio. Eu preciso saber se eu trato os dados pessoais, onde eles estão e por onde passam para eu poder redobrar a segurança desses repositórios e sistemas. Aquilo que gera um desafio gigantesco para o time de segurança ficou ainda maior.

Bom, só fica mais complexo. A LGPD veio para deixar bem claro que os donos dos dados pessoais são os próprios titulares. Basicamente, todas as empresas que tratam esses dados têm a obrigação de prestar contas deles aos seus donos. Então, nada de letrinha miúda, um link escondido em algum lugar no site, um sistema que pede ainda mais dados para poder funcionar. A experiência de pedir um relatório mostrando todos os dados que a empresa tem sobre você tem que ser superfluida. Fácil de falar, claro. Agora, pensa como você vai [pegar] todos os dados de todos os sistemas e repositórios de uma empresa, reunir tudo isso para uma determinada pessoa e depois emitir um relatório informando para ela os resultados. A gente viu a questão da segurança. Essa aqui é a questão dos titulares. É mais uma face desse desafio.

E, para completar o quadro, tem a própria ANPD. A gente falou de segurança, que é uma situação atípica, ou deveria ser. Os titulares vão ser esporádicos, mas tem o *business* da empresa, tem o dia a dia. E é aqui que entra o terceiro item. Lembra dessa analogia de que dados pessoais são radioativos? Você precisa exercer uma atividade antes de mais nada que justifique o tratamento desses dados. Não é qualquer empresa que faz qualquer negócio, qualquer *business*, que pode tratar dados pessoais. E, a qualquer momento, você pode ser solicitado a apresentar o ROPA, esse Record of Processing Activities, que é um registro que vai listar todos os dados pessoais e sensíveis tratados pela empresa. [Ele] vai falar das origens desses dados, das atividades em que eles são utilizados, quem são os consumidores, a finalidade, a base legal que sustenta toda essa operação — que é uma defesa mesmo, é uma tese jurídica que defende o seu negócio —, a data em que esses dados vão ser descartados, só para nomear algumas informações. O ROPA é gigante, isso aqui é só um pedacinho dele.

Eu acho que isso dá um bom panorama, pensando no desafio. Infelizmente, a LGPD acabou virando um veículo de propaganda e marketing para um monte de empresas. Faça uma pesquisa rapidinho no Google de adequação à LGPD e aparece um monte de consultoria milagrosa e sistemas que fazem a adequação da sua empresa em uma semana. Se pagar à vista no Pix ainda traz o seu amor de volta! Então, a LGPD é tipo uma prova de Iron Man. Você pode contratar um treinador, uma academia, equipamento de última geração, suplemento e tal, mas no final do dia, é você que tem que nadar, pedalar e correr. É uma jornada da empresa, não é um conjunto de itens que a gente vai dando “check” e em algum momento termina. Inclusive, conecta diretamente com a alfabetização de dados de que a gente falou, porque depende muito de aculturar todo mundo nesse tema. Para dar um exemplo bem besta, só o conceito de dados pessoais parece óbvio, fácil de imaginar, e já embola uma galera. Dados sensíveis, então, lascou! Alguém precisa olhar para aquele mar de dados e apontar o que é crítico, o que deve ser avaliado imediatamente, para fazer uma avaliação de impacto sobre a proteção de dados, o famoso DPIA, que é um documento jurídico bem técnico, bem denso. Alguém tem que fazer uma avaliação de legítimo interesse — a gente faz uma avaliação para ver se você realmente pode fazer aquele tratamento de dados — e, por fim, desenvolver, desenhar todo esse ROPA, que é esse registro vivo de todas as atividades que a empresa faz

que envolvem dados pessoais. Então, sério gente, eu não consigo imaginar uma jornada de adequação à LGPD sem um time jurídico especializado em proteção e privacidade de dados, seja interno ou *as-a-service*, como a gente fala, uma consultoria. Ele é obrigatório.

Mas nem tudo são trevas. Claro que existe uma série de ferramentas que vão tornar essa jornada menos desgastante e impactante para todos. Na frente da segurança, por exemplo, só o Google Workspace fez. Não estou nem falando de empresas com um modelo de negócio superavançado e processamento de dado pesado. Vamos pensar só no dia a dia, que geralmente é ali que a gente negligencia um pouco as coisas. Só no dia a dia, o Google, por exemplo, o Google Workspace já vai ajudar um monte! Primeiro, porque ele cria um espaço de trabalho que não é só seguro, mas acima de tudo, é todo catalogado. Tudo, tudo, tudo o que acontece ali está sendo catalogado. E-mail, Agenda, Chat, Meet, todos os Documentos, as Planilhas, Apresentações, os Formulários, Sites, tudo está protegido por padrão. Se tiver o segundo fator de autenticação, então, mesmo que alguém descubra a senha, não vai conseguir acessar. Ou seja, para vazarem uma planilha — a gente falou do vazamento de 220 milhões e tal — para vazarem uma planilha com dados, por exemplo, de funcionários da empresa, tem que querer muito, se você estiver com o segundo fator de autenticação ativado. E o mesmo serve para Cloud. Ela já vem com um monte de sistemas de segurança por padrão, como o IAM, que vai fazer o controle superfino de acesso a todos os recursos da Cloud, além de toda a segurança que já está embutida na infraestrutura.

Então, para a outra parte do desafio — que é quando a gente vai pensar em processamento de dados mesmo — tem o Google Data Catalog, de que a gente já tinha falado lá em cima, e aqui ele se agiganta, porque pode ser utilizado dentro do coração da gestão da privacidade, justamente para montar esse ROPA. Ele tem a capacidade de varrer os repositórios de dados em busca de dados pessoais até sensíveis, permite uma visão centralizada de tudo isso sem que você tenha que ficar movendo dados para lá e para cá.

Então já dá para responder perguntas, como quais são todos os dados pessoais que eu tenho. O IAM mostra quem tem acesso a esses dados e por aí vai. Claro que a gente está falando sobre GCP. Existem algumas soluções no mercado já prontas para isso também, mas para quem, de repente, não tem uma complexidade tão grande assim, só com aquilo que já tem pronto dentro do GCP, já dá para montar um baita de um sistema de gerenciamento de dados pessoais e se organizar para a ANPD!



**Marcelo Gomes:** Que legal, Michel! Eu acho bacana porque você trouxe vários tópicos que estão superlatentes e superimportantes. Você comentou sobre o segundo fator de autenticação. A gente sabe que, hoje, ao usar simplesmente uma senha, a gente não está mais seguro, visto a facilidade que existe em uma quebra de senha. Então usar esses contextos e realmente utilizar o que há de serviços e soluções para conseguir auxiliar na proteção é muito importante.

Eu acho que, até agora, a gente falou bastante sobre por que os líderes investem em todo o cenário de dados, sobre o contexto em volta dos dados e essas soluções de análise de dados para as empresas poderem ter uma

estratégia mais assertiva. Mas acho que a gente pode falar um pouquinho mais sobre como isso funciona na prática. Afinal, todos os líderes estão à frente de equipes e tem times de diversos setores dentro das empresas em si. Tem time da área financeira, de marketing, comercial, enfim, área de tecnologia. Isso nos leva à pergunta: como adotar uma cultura *data-driven* dentro da empresa para que todos se beneficiem dos dados e entendam a importância dessas análises? Até levando em conta todos os contextos de que você falou, de complexidade para a gente construir essa linhagem do dado como um todo.



**Michel Avila:** Show! Aqui, como você falou, a gente não tem como não revisitar. A gente revisita a alfabetização de dados, que é obviamente um pré-requisito para esse movimento todo acontecer. Então não tem como a gente falar de uma... não tem como a gente chegar numa empresa que usa dados, que busca a organização e utilização de dados para tomar decisões, sendo que essa empresa não passou, ou passa constantemente, por uma alfabetização de dados. Mas é um processo evolutivo. Uma vez que esses conceitos básicos já estão circulando pela empresa e a turma está nivelada, se for uma empresa — e vou bater bastante nessa tecla — se for uma empresa que tem um clima, uma cultura fértil, vai rolar organicamente. Você coloca em curso esse tipo de treinamento, esse tipo de *mindset* e as iniciativas começam a pipocar aqui e ali, e você só tem que ir coordenando. É crucial que exista um lugar onde tudo se encontra, onde todos, de alguma forma, podem ter visibilidade desses trabalhos e trocar experiências, emprestar os expertises e progredir. Porque, primeiro: cada time tem desafios diferentes. Nem todos os times vão ter todos os *skills* que são necessários para de repente coletar um dado. Um time, de repente, um pouco mais focado no comercial, um time do departamento pessoal. Então, vai faltar *skill* e, se essa galera tem um lugar para conversar e para trocar figurinha, pode ser que ali a gente consiga servir, os grupos e as áreas possam se servir umas das outras para conseguir montar essa visão que eles querem, organizar esses dados, fazer experimentos.

Eu vejo isso tudo como se fosse um grande *hackathon*, sabe? Bem multidisciplinar e bem experimental mesmo. Se não houver na empresa um *skill* necessário — porque às vezes falta, às vezes é uma empresa que não tem aquilo ali, não é do cotidiano dela — vale tudo! Contrata, terceiriza, busca alguma opção *open-source*, consultoria. O mais importante aqui é dar esse espaço para que esse movimento aconteça, porque isso vai mandar uma mensagem clara de que essa busca por formas de direcionar o dia a dia da empresa usando dados é o único caminho. Como diz o Mandalorian, “this is the way”. Então esse ambiente que se cria, é ele que vai dar esse tom. Depois você vai trabalhar um pouco mais com a camada de gestão para coordenar as atividades. De repente, você vai ter que priorizar alguma atividade dessas em detrimento de outra. Vai muito da maturidade da empresa, mas tendo as condições, tendo uma cultura, tendo um clima fértil para isso, essas coisas vão acontecer automaticamente. Elas vão pipocar, porque a galera vai sentir falta de fazer isso, vai sentir falta de ter uma previsibilidade, vai sentir falta de olhar para o seu resultado, de repente, de pessoas no final do ano e ver: o que aconteceu, como é que eu posso medir, como é que eu posso ver se as minhas contratações estão sendo melhores, como é que eu posso ver se o clima e a cultura da empresa estão melhorando ou não, como é que eu posso ver se meu cliente vai dar um *churning*, será que não dá para fazer um modelo preditivo de *churning*? Tem “n” coisas que dá para fazer. E, na prática, não são coisas

complexas, sabe? O conceito em torno, se quiser transformar isso numa máquina, pode ficar mais complexo, se quiser escalar um treco desses. Mas até chegar a esse momento, a galera já roda bastante. A galera roda, de boa, com planilha Excel, assim, se vira! Então, num primeiro momento, se a empresa não tem isso, eu focaria muito mais numa parte mais inicial, de resolver alguns problemas com as ferramentas que tem, com o *skill* que já tem dentro da empresa, focando muito mais no *mindset*, sabe? Para a galera dar esse primeiro passo, sentir um gostinho de ter dados para tomar as decisões e dar o tempo. Você precisa de tempo para medir, tempo para ver os resultados, tempo para dar um feedback para aquele sistema e a galera olhar e falar: “Cara, super curti! Funcionou! Eu quero melhorar. Como eu faço agora?”. Então seria mais ou menos nessa linha.



**Daniel Leite:** Respostas extremamente completas. Eu acho que toda a argumentação que você colocou aqui faz total sentido. Depende muito também da maturidade, do momento de cada empresa. Como você comentou agora, às vezes, trabalhar talvez numa planilha inicialmente possa ser um caminho, pelo menos para entender qual é a sequência. Mas, claro, esse nível de maturidade vai permear todas as fases que você comentou: o entendimento do dado, que direcionamento pode ser dado a ele, qual a transformação, como ele pode ser melhor consumido. Foi uma aula hoje, Michel! Muito obrigado!



**Marcelo Gomes:** Sensacional, Michel! Eu faço minhas as palavras que o Dani acabou de falar para você. Foi uma aula. A gente pôde aprender bastante contigo, revisar vários temas que são superlatentes, superimportantes, que a gente está vivendo. E acho que, na nossa entrevista, a gente vai ficando por aqui. A gente queria agradecer demais a sua participação e por ter compartilhado todos esses temas, certo, Dani?



**Daniel Leite:** Isso mesmo, Marcelinho. Michel, muito obrigado! Foi extremamente importante, muito prazerosa sua participação aqui conosco. Ficamos muito felizes e com certeza os nossos ouvintes também ficaram, e ficarão ouvindo esse podcast sem parar, com certeza.

**Vinheta:** De cabeça na nuvem.



**Marcelo Gomes:** Você sabia que a quantidade de dados armazenados na nuvem é praticamente inimaginável? Pois é, até o ano de 2025, é provável que existam mais de 100 ZB de dados na nuvem. Para você ter uma noção básica sobre isso, um zettabyte é igual a 1 trilhão de gigabytes. Ou seja, basicamente, haverá mais informações armazenadas na cloud do que fora dela. Estima-se um total de 175 ZB até 2025. Por isso que as questões sobre segurança digital são tão importantes. A gente fala muito sobre isso porque esse é realmente o futuro de armazenamento dos dados.



**Daniel Leite:** E, falando nisso, você consegue imaginar quantas pessoas usam serviços de armazenamento em nuvem? Pois é, esse é um cálculo muito difícil de ser feito, mas dá para imaginar um caminho para isso. O Google Drive é o serviço mais popular da internet. Entre todos os usuários de serviços em nuvem, 94,4% usam o Google Drive, que chegou a uma marca histórica de 1 bilhão de usuários. Em 2020, a plataforma como um todo do Google Workspace atingiu 2 bilhões de usuários. É muita gente!

Obrigado a todos que estão nos ouvindo, não esquecendo do nosso e-mail [googlecloudcast@google.com](mailto:googlecloudcast@google.com). Qualquer dúvida, qualquer tipo de interação, estamos aqui para ouvi-los, ler as suas mensagens e depois responder com todo o carinho. Michel, obrigado mais uma vez.



**Michel Avila:** Boa, Dani! Muito obrigado, gente, Dani, Marcelo, pela oportunidade de compartilhar esse episódio com vocês e trocar um pouco de ideia. Foi um privilégio realmente poder participar! Muito obrigado!.



**Daniel Leite:** Maravilha! Ficamos aqui com mais um episódio do Google Cloud Cast. Até a próxima!



**Marcelo Gomes:** Obrigado, gente!