



CONSEGUIREMOS AGUENTAR A VERDADE?

Tiago Santos

Ao trazermos à tona nossas entranhas econômicas nas relações cliente X mídia X tecnologia X jornalismo, teremos um processo doloroso pela frente

O escândalo da Cambridge Analytica está de volta. Dessa vez, sob os holofotes do documentário da Netflix, Privacidade Hackeada ou The Great Hack, em inglês. Mas se intenção era aprofundar o caso da participação da empresa em processos eleitorais pelo mundo, mais especificamente nos casos do Brexit, na Inglaterra, e na eleição de Donald Trump, nos Estados Unidos, a produção peca ao não falar de um dos principais pontos referente aos processos de big data, inteligência artificial e data science, ainda que demonstre como dados podem ser minerados e utilizados para ações de marketing digital. Então, antes de comentar especificamente sobre o caso da Cambridge Analytica e do documentário, vale detalhar um pouco sobre esses processos e como são utilizados no mundo atual.

Jeff Hawkins, um dos criadores do Palm (precursor dos smartphones) e um dos “tech” neurocientistas que mais tem contribuído com mercado de psicologia tecnológica nas últimas décadas, possui uma excelente definição do que é inteligência: “Algoritmo de aprendizado cortical”. Ele diz que nosso cérebro é um sistema de distribuição sequencial de memória e que o aprendizado e a inteligência ocorrem por meio da criação de modelos semânticos de sequências de dados captados. De forma mais simples, ele acredita que gravamos dados no nosso cérebro em uma determinada ordem sequencial, que gere significado. Ao encontrarmos dados A, B, C, D... os gravamos em uma ordem, gerando um significado específico. Ao reencontrarmos parte deles no mundo (B, D...), podemos prever os dados futuros, ou inferir nos dados passados, de acordo com o histórico e o significado que registramos. Definiu-se, então, inteligência artificial (IA) como a capacidade de um computador de reproduzir esse processo:



identificar dados e, baseado nas sequências pré-registradas, tomar uma decisão de inferência anterior, ou previsão futura.

No mundo atual, a quantidade de dados registrados e processados pelas empresas se tornou tão grande que foi necessário estabelecer tecnologias e estratégias específicas para lidar com eles. Diferentemente do abecedário, que tem a quantidade de dados limitada, em uma empresa esses dados são massivos. A estruturação desses dados para um melhor aproveitamento e automação deles é o que chamamos de big data.

Porém, vamos lembrar que o processo de inteligência não se resume a gravar e processar dados, mas em criar um significado para eles. Para isso, é necessário encontrar sequências lógicas em dados que não são tão estruturados quanto A, B, C, D, E... imagine os dados de alguém que está constantemente em viagem. Como encontrar uma lógica entre essa pessoa consumir um cachorro-quente no aeroporto de São Paulo, horas depois acessar um site no celular de Porto Alegre e terminar o dia comprando produtos de limpeza por um aplicativo de entrega em São Paulo. Mais do que isso, imagine como encontrar semelhanças entre esse comportamento e o de outras pessoas e assim tomar decisões empresariais. Isso é data science.

Depois dessa digressão, vamos voltar ao caso da Cambridge Analytica, que se posicionava como uma empresa de data science, ou seja, analisava perfis, segmentava a audiência online e sugeria o tipo de ação (campanhas de mídia) a ser tomada com aquele público. Os dados são em sua maioria gravados em empresas como Facebook, Google, Adobe, Oracle, IBM, Salesforce. E são prioritariamente ativados (ou seja, a mídia é comprada com base nesses dados) no Facebook e Google. Mas o Facebook não é a origem dos dados (ao menos, não a única), e Alexander Xis, CEO da Cambridge Analytica, afirma isso em sua audiência em 2018. Ex-funcionários da empresa também afirmaram que a plataforma era utilizada para compra de mídia e que a empresa tinha no seu centro a compra e ativação de mídia digital. Então, as grandes perguntas que



ficam são: “Quem gera a captação desses dados?”; “Quem se beneficia desse sistema?”.

A resposta é dolorosa. Grande parte da captação desses dados é feita nos sites e plataformas dos que criticam ou expuseram o escândalo, como The Guardian. Sim, o site do mesmo jornal responsável por revelar a situação da Cambridge Analytica. E isso pode ser comprovado em uma rápida análise de um datamapping, mapeamento que identifica a origem da captação e distribuição de dados de qualquer empresa que esteja na internet. No caso do The Guardian, é possível notar a quantidade imensa de empresas que capturam dados fornecidos por ele. Ao todo, são mais de 80 diferentes empresas capturando dados de leitores do The Guardian. Essa captura ocorre por códigos incluídos no site pelo próprio jornal, com o objetivo de melhorar a entrega publicitária dentro do site. Dados que variam de informações técnicas do cliente que está acessando o site (como tipo de dispositivo, tamanho de tela, sistema operacional), passam por dados de comportamento online (quais páginas estão sendo acessadas, notícias compartilhadas, vídeos assistidos) e chegam a dados mais individuais, como geolocalização.

Mas, por que fazem isso? Por que uma empresa como The Guardian captura e compartilha os dados de clientes dessa forma? Porque o real produto de uma empresa de conteúdo é o conteúdo! Elas sobrevivem por meio de receitas de publicidade, que, por sua vez, necessitam desses dados dos consumidores para traçarem suas estratégias de mídia. A estrutura de funcionamento o mercado sabe: com o avanço as tecnologias digitais, a publicidade vem se tornando cada vez mais direcionada; as empresas buscam atingir o cliente certo no momento exato, com o produto mais adequado para ele. E como fazem isso? Exatamente da mesma forma que a Cambridge Analytica fez: utilizam dados capturados em sites de conteúdo e plataformas de serviços sociais para identificar perfis de consumidores. Segmentam esses clientes e determinam jornadas de comunicação para cada um desses segmentos e, a partir daí, entram em um leilão com outras empresas para a compra de espaço publicitário nos mesmos sites que geraram os dados para atingir esse cliente. Os sites ficam com uma



parte desse valor pago, as plataformas de mídia e segmentações (Google, Facebook, Criteo, TailTarget), com outra.

Legislações como a europeia (GDPR) e a brasileira (LGPD) estão buscando regulamentar essa captura e utilização de dados. O próprio The Guardian está se adaptando e permitindo que o consumidor defina se quer ou não compartilhar os seus dados. É um avanço importante, mas é importante a consciência de que somos muito mais responsáveis por acontecimentos como o da Cambridge Analytica do que queremos acreditar. Quando definimos que iremos consumir o trabalho de alguém gratuitamente, obrigamos essas empresas a buscar alternativas de receita. É a velha história: não existe almoço grátis.

E isso fica claro no documentário Privacidade Hackeada e em três de seus personagens: Brittany Kaiser, ex-funcionária da Cambridge Analytica; David Carroll, professor americano que abriu um processo pra recuperar seus dados; e Carole Cadwalladr, jornalista do The Guardian que revelou o caso. Eles resumem o mercado atual: o consumidor que não quer pagar pelo serviço que consome e acaba sendo transformado em produto, sem saber ou fingindo não saber o custo da gratuidade; a empresa ávida por anunciar e nem sempre com noção do caminho para obtenção dos dados dos consumidores; e a empresa ou plataforma de conteúdo que precisa ser gratuita, e para isso transforma seus consumidores em um produtos de venda para anunciantes.

Nesse thriller jurídico-tecnológico que iremos viver nos próximos anos, a ponta solta é reconhecer a nossa responsabilidade como consumidores e das empresas de mídia como captadores e vendedores desta informação. Ao trazermos à tona nossas entranhas econômicas nas relações cliente X mídia X tecnologia X jornalismo, teremos um processo doloroso pela frente. Conseguiremos aguentar a verdade?