IA no RH: decisão sobre pessoas exige revisão humana e trilha auditável

Em um dos maiores bancos do país, a lista de candidatos que chega ao recrutador já vem ordenada por um sistema de IA que ranqueia os currículos por aderência à vaga

Rodrigo Gomes (*)



Worawut Prasuwan_CANVA

Empresas já usam a IA para responder à dúvida de funcionários, e até mesmo para avaliar a coerência das avaliações de desempenho feitas pelos gestores. Mesmo que a decisão final continue com uma pessoa, o RH precisa conseguir dizer, quando alguém pergunta, por que o sistema recomendou aquilo e quem decidiu seguir a recomendação.

E essa distância entre usar e explicar aparece nos números de adoção da própria tecnologia. O relatório 2026 Global Human Capital Trends, publicado pela Deloitte em março de 2026 a partir de pesquisa com mais de 9 mil líderes de negócios e de RH em 89 países, descreve a IA incorporada à contratação, ao desempenho e às decisões cotidianas: 60% dos executivos usam a tecnologia na tomada de decisão, enquanto 5% afirmam gerenciar bem esse uso, e 42% dos trabalhadores dizem que suas organizações não avaliam o impacto da IA sobre as pessoas.

No Brasil, a pesquisa Confiança, Atitudes e Uso de Inteligência Artificial, conduzida pela Universidade de Melbourne com a KPMG e publicada em 2025, com mais de 48 mil respondentes em 47 países e dados coletados entre novembro de 2024 e janeiro de 2025, mostra que 86% dos respondentes brasileiros afirmam que suas empresas já usam IA, enquanto 55% se dizem dispostos a confiar nas decisões tomadas pela tecnologia e 49% admitem ter cometido erros por confiar em resultados que ela entregou.

Quando quase metade dos respondentes admite ter errado por seguir o sistema, a revisão humana passa a ser o mecanismo pelo qual a empresa protege a própria decisão, antes de qualquer exigência externa.

Essa exigência externa, aliás, já apareceu no Marco Legal da Inteligência Artificial, o PL 2338/2023, aprovado pelo Senado em dezembro de 2024 e que está à espera do parecer do relator, com uma ementa que fala em uso ético e responsável da IA com base na centralidade da pessoa humana.

Revisão humana: em qual ponto a decisão passa do agente ao gestor

Levada para a rotina do RH, essa centralidade se traduz em uma linha que separa o que um agente executa de ponta a ponta do que precisa parar em uma pessoa, e a linha é traçada pelo efeito da decisão sobre quem é afetado, o que explica por que a mesma ferramenta pode operar sozinha em um processo e apenas preparar o terreno em outro.

Conferir se a documentação de admissão está completa, responder a uma dúvida sobre saldo de férias, apontar uma marcação de ponto sem justificativa, agendar entrevistas e resumir currículos são tarefas nas quais o erro é reversível e barato, e o agente pode concluí-las com registro. Cortar candidatos de uma lista, atribuir nota de desempenho, recomendar promoção, sugerir desligamento ou aplicar sanção são decisões nas quais o erro atinge a carreira de alguém, e nesse ponto o sistema entrega uma recomendação com motivos e a pessoa decide.

Revisão humana só cumpre o papel quando o revisor conhece os motivos da recomendação, dispõe de tempo para examiná-los e tem autoridade para discordar, porque sem essas condições, o RH vai depender somente da correlação realizada por um algoritmo.

Na prática, isso pede que cada fluxo de RH tenha um ponto de parada definido antes do agente entrar em operação, com o nome do responsável pela decisão e o formato no qual a recomendação chega a ele, com os critérios que pesaram e a confiança do sistema naquele resultado.

Quando a recomendação chega como um sim ou não sem contexto, o revisor tende a seguir o sistema, e o registro posterior mostrará que a decisão foi, na prática, automatizada. Quando chega com os motivos, o revisor pode discordar, e a discordância vira dado para corrigir o modelo, o que devolve à empresa o aprendizado que a automação sozinha não produz.

Trilha auditável para sustentar a decisão humana

Explicar a decisão e permitir que seja contestada, como o texto do Senado exige para os usos de alto risco, só tem efeito prático se a empresa consegue reconstruir a decisão depois de tomada, e é isso que as medidas de governança previstas para quem aplica esses sistemas descrevem ao pedir documentação ao longo do ciclo de vida, avaliação da acurácia e da robustez dos resultados, apuração de resultados discriminatórios, registro do grau de supervisão humana que contribuiu para cada resultado e informações que permitam interpretar o funcionamento do sistema.

Traduzida para o dia a dia, a trilha de uma decisão assistida por IA precisa responder a perguntas diretas: quais dados o agente viu, qual versão de modelo, regra ou instrução estava em uso, qual foi a recomendação e com quais motivos, quem revisou, o que decidiu e por que, se houve contestação e qual foi o desfecho. Cada resposta corresponde a um registro com data, e o conjunto precisa ter prazo de guarda e controle de acesso definidos, porque envolve informação sobre pessoas.

A esse núcleo se soma um teste periódico dos resultados por grupo, para verificar se a triagem ou a avaliação está excluindo mulheres, pessoas acima de determinada idade ou candidatos de certa origem em proporção diferente da entrada, apuração que o próprio texto legal pede a quem aplica o sistema.

Montada assim, a trilha serve ao colaborador ou candidato, que recebe a explicação que tem direito de pedir, ao gestor, que ganha a demonstração de que houve revisão humana com conteúdo quando a decisão é questionada dentro da empresa ou em juízo, e à própria organização, para a qual a discordância registrada entre recomendação e decisão final é o indicador mais direto de que o modelo precisa de ajuste. Nada disso depende de esperar a regulamentação, já que o texto do Senado remete detalhes a normas posteriores, mas os elementos da trilha decorrem da própria lógica de decidir sobre pessoas com apoio de um sistema.

O critério que fica para o decisor independe do texto final que sair da Câmara: quanto mais etapas o agente assume, mais valor tem o ponto no qual a decisão volta para uma pessoa com os motivos à mão e autoridade para discordar, e mais barato é construir esse ponto antes que a primeira contestação obrigue a reconstruí-lo.

(*) Head da unidade de negócios Process Solutions da Selbetti.

ESPAÇO JURÍDICO



Giovanna Araujo (*)

O regulador de IA que ninguém elegeu

Lula foi reeleito com 51,79% dos votos contra 48,21% de Flávio Bolsonaro, numa eleição em que a abstenção caiu em relação ao primeiro turno. O placar, porém, não encerrou a disputa — mudou seu endereço. Na noite da vitória, o derrotado anunciou que pedirá ao TSE a anulação do pleito, por causa da decisão do ministro Cristiano Zanin, do STF, que prorrogou a votação até as 19h diante da redução da frota de ônibus em ao menos dez estados. "Não vou aceitar que a regra do jogo tenha sido mudada em cima do jogo", disse. Do outro lado, a campanha vencedora mantém viva a ação que pede a cassação da chapa adversária por "manipulação algorítmica" da Meta. Os dois finalistas contestam a mesma eleição, por razões opostas, no mesmo tribunal.

O pedido de anulação não menciona inteligência artificial. Mas aterrissa num tribunal que passou o ano arbitrando IA: julgou o primeiro deepfake de campanha, impôs isonomia algorítmica à Meta sob multa diária, instalou uma sala de riscos digitais com as plataformas dentro de casa. O próprio presidente do TSE, Nunes Marques, criticou a prorrogação falando em "hiperjudicialização" — embora afirme que ela não alterou o resultado. A palavra diz mais do que parece: a judicialização não começou no domingo; foi o método do ciclo inteiro.

O balanço empírico saiu na sexta anterior ao segundo turno. O monitoramento VigIA (Lupa/Unicamp) confirmou 1.423 conteúdos políticos criados ou alterados por IA — quase 21 por dia. Metade era deepfake; 768 circularam sem o rótulo que a Justiça Eleitoral exige, a maioria nas redes da Meta. E o dado que desmonta a tese do infrator anônimo: perfis oficiais de candidatos (548) quase empataram com usuários comuns (560) entre os autores. "A rotulagem falhou", resumiu Cristina Tardáguila, da Lupa. A régua existia; quem disputava o poder descumpriu quase tanto quanto quem se esconde atrás de conta falsa.

A tensão central das democracias é velocidade versus legitimidade. O Brasil fez essa escolha sem debatê-la: com o PL 2338 parado na Câmara desde novembro de 2025, quem regulou a IA da eleição foi o TSE — por resolução, liminar e monitoramento. Funcionou onde havia CNPJ para intimar. Mas regra escrita por juiz carrega o rosto do juiz: quando a decisão aperta, discute-se a suspeição do ministro, não o mérito da norma. Lei tem a legitimidade de quem a votou. Resolução tem a de quem a assinou — e essa se contesta no grito.

A desculpa, agora, acabou. A votação do PL 2338 foi adiada "para depois das eleições"; as eleições passaram e deixaram o relatório de impacto pronto: os números do VigIA, a jurisprudência fragmentada construída caso a caso, os relatórios de conformidade que as plataformas devem entregar ao TSE até dezembro. O Congresso tem em mãos o dossiê que dizia esperar.

Para as empresas, a lição não é eleitoral. Na ausência de lei, reguladores de fato — tribunais, agências, juízes de plantão — escrevem a regra no caso concreto, e quem opera IA descobre a régua quando já é réu. Rotulagem de conteúdo sintético, trilha de auditoria, resposta documentada em horas: o padrão que o TSE impôs às plataformas é o ensaio geral do que vem para todos os setores. Quem esperar a lei para se organizar vai ser regulado por liminar.

(*) - Sócia do Araujo Advs - advogada especializada em Governança de Inteligência Artificial

A IA virou a melhor contadora de histórias do mercado financeiro

Dionathan Henchel (*)

A conversa sobre IA nas finanças costuma girar em torno de precisão, velocidade e custo. Qual modelo erra menos, qual automatiza mais, qual entrega mais rápido. Essa é a discussão que todo mundo enxerga. Depois de concluir o programa Artificial Intelligence for Financial Services, no MIT Sloan, volto com uma preocupação que, na minha avaliação, vai pesar nos próximos anos: a IA já escreve análises que o mercado prefere ler. E preferir não é o mesmo que acertar.

O programa reúne professores que ajudam a moldar a forma como pensamos mercado, dados e IA. Entre eles, Andrew Lo, que propôs olhar os mercados pela lente da biologia evolutiva, como organismos que se adaptam, e que hoje estuda como a IA se encaixa nessa visão. Há também especialistas em construção de agentes, em governança e privacidade de dados, em interpretabilidade de modelos e em negociação que debateram as condições em que uma instituição financeira pode confiar no que a IA entrega.

Um dos achados mais fortes do curso foi entender como decidimos e que a razão não opera sozinha. Decisões dependem da emoção, e é por isso que, em crises como 2008 e a pandemia, investidores migraram para a liquidez justamente no pior momento. Quando o futuro é incerto, o cérebro busca histórias com causa e efeito, porque elas devolvem uma sensação de controle. Essa preferência por narrativas é uma vantagem evolutiva, e é também o que torna o mercado tão sensível a quem conta bem a história.

Os grandes modelos de linguagem são, no fundo, máquinas de narrativa. Avançam uma palavra por vez, minimizando a surpresa e a cada nova geração ficam melhores em soar como um conselheiro, acolhem primeiro e explicam depois com tom seguro. Num teste cego conduzido no curso, 160 leitores deram nota maior a relatórios de análise gerados por IA do que a relatórios reais de analistas de mercado. Acolhimento na linguagem, no entanto, não equivale a compreender emoções nem garante a qualidade do conselho.

É aí que mora o risco. Uma resposta errada da IA soa exatamente como uma certa. Em testes com 100 mil notícias financeiras, as previsões em que o modelo tinha alta confiança interna acertaram cerca de 65%, enquanto as de baixa confiança ficaram perto de 51%, praticamente um cara ou coroa. Já a confiança que o próprio modelo declara sobre si mesmo se mostrou pouco confiável. Com técnicas de interpretabilidade, pesquisadores observaram que o que o modelo diz ter levado em conta nem sempre é o que de fato pesou na decisão. O modelo não paga custo por errar nem guarda memória de ter errado. Ele só continua falando bem.

Isso muda o que significa usar IA com responsabilidade. Se a fluência não é prova de qualidade, ela precisa ser tratada como um risco a ser medido ao invés de ser um sinal de confiança. Logo, exige olhar para os sinais internos do modelo, e não para o tom da resposta, auditar em que ele está de fato apostando e posicionar a revisão humana onde o erro custa mais e onde as pessoas têm vantagem. E um aval humano mal calibrado pode fazer uma resposta errada parecer segura.

A disputa da IA nas finanças continuará sendo travada em capacidade e em velocidade. Mas num mercado movido por medo e por histórias, a IA tem uma vantagem que ainda não aprendemos a medir. Convencer ficou barato, acertar continua caro e quem aprender primeiro a desconfiar da fluência vai errar menos.

(*) Sócio-fundador da Vertrau.