Durante 2023, a Inteligência Artificial (IA ou AI, do termo em inglês Artificial Inteligence) emergiu como o principal tema tecnológico, ampliando seu alcance para além das grandes corporações. Versões comerciais estão sendo desenvolvidas para uma variedade de usos, incluindo pesquisa, produção de conteúdo, planejamento e automação de processos, oferecendo benefícios significativos em diversas áreas profissionais.

A Inteligência Artificial (IA) é uma ferramenta na qual um humano emite um comando por meio de uma caixa de texto, conhecida como prompt de comando ou CMD. Esse comando é processado por uma plataforma ou sistema computacional programado para fornecer respostas, informações e outros conteúdos relevantes ao questionamento humano.

Por meio de mecanismos padrão de "machine learning", a IA utiliza algoritmos computacionais para buscar respostas e padrões na Internet. Esses algoritmos ensinam a máquina a identificar conteúdos relevantes com base no número de sites que contêm palavras ou imagens-chave semelhantes às solicitadas, levando em consideração sua relevância e popularidade. Após analisar, coletar e compilar as informações, a resposta é formatada conforme solicitado e apresentada automaticamente, frequentemente em questão de segundos.

Nos primeiros meses de lançamento comercial deste tipo de ferramenta, tanto o mercado quanto os desenvolvedores constataram que os resultados não se limitaram a benefícios e agilidade. Houve casos em que respostas sensíveis ou que poderiam gerar riscos físicos foram apresentadas pela ferramenta sem supervisão ou intervenção humana adequada, possibilitando o uso inadequado das ferramentas.

Mas, afinal, o que é AI e como ele funciona? Pode-se dizer, grosso modo, que seria uma ferramenta através da qual um comando é dado por um humano a uma caixa de texto, especificando-se à ferramenta o que se deseja. Tecnicamente, e até mesmo para que você possa pesquisar e se aprofundar futuramente, isso é chamado de um prompt de comando ou CMD. Seguindo, este texto é então "lido" e processado pela ferramenta – uma plataforma ou sistema operacional/computacional – que foi programada para oferecer respostas, informações, imagens, apresentações, esquemas, fluxos, etc. relacionadas aos termos do questionamento feito pelo humano.

Como isso é feito por trás das cortinas? Através de mecanismos padrão de "*machine learning*", algoritmo computacional que "ensina" a máquina a buscar respostas e padrões na Internet com base no número de sites contendo palavras ou imagens chaves semelhantes às solicitadas, de acordo com a sua relevância e número de vezes em que aquele determinado conteúdo é linkado a outros conteúdos. Feita esta análise, coletada e compilada a

informação e formatado o resultado na forma requerida por aquele que consulta, a resposta é apresentada automaticamente, muitas vezes em poucos segundos.

É oportuno destacar que, logo nos primeiros meses de lançamento comercial deste tipo de ferramenta, o mercado e os próprios desenvolvedores verificaram que os resultados não foram apenas benefícios e agilidade. Respostas eram apresentadas pela ferramenta sem nenhuma supervisão ou intervenção humana, ao solicitante, de modo que temas sensíveis, ou que poderiam gerar risco físico, inclusive, eram respondidos sem uma curadoria adequada, abrindo margem para a utilização indevida das ferramentas.

Assim, apesar de termos uma ferramenta tão inovadora e com potencial real de auxiliar e minimizar o tempo de trabalho de tantos outros profissionais ou ofertar experiências muito mais otimizadas a usuários de sites comerciais, verificou-se também a necessidade de debate e aprofundamento de sua utilização, ou mesmo de uma regulação adequada.

Neste sentido, o machine learning, nesta época era uma técnica capaz de ensinar softwares a buscarem respostas de acordo com técnicas pré-programadas ou baseadas em estatísticas previsíveis de conexões entre palavras-chave. Em última instancia, ferramentas de IA só seriam capazes de operar com situações previsíveis, baseadas em ocasiões ou hipóteses já existentes, de forma automática, sem fazer qualquer juízo de valor ou aplicar técnicas de bom senso (características impossíveis de se aplicar a máquinas pelo menos até aquele momento). Note-se, assim, que somente humanos são capazes de usar bom senso e fazer juízos de valor. Isso tornava as ferramentas de IA incapazes de reagir satisfatoriamente (ou até mesmo corretamente) a determinadas situações imprevisíveis e inusitadas. Leve-se também em consideração que as ferramentas de IA eram incapazes de criar, sozinhas, padrões novos de machine learning, sendo apenas capaz de executar os padrões já embarcados em sua codificação. Necessitariam para tanto de constante atualizações de seus códigos e algoritmos, mas sempre de forma reativa, nunca proativa ou criativa.

E qual a consequência jurídica desta condição? Pois bem, sem entrar especificamente em qualquer dispositivo específico do Direito, mas já a partir de uma visão bastante holística da Teoria Geral do Direito, existem boas chances das ferramentas de IA responderem a determinados questionamentos inesperados, imprevistos ou delicados e sensíveis o suficiente que demandem a aplicação do juízo de valor e do bom senso, definitivamente de maneira imoral, antiética e ilegal.

Imagine-se fazer uma consulta sobre a atual Guerra da Palestina. Dependendo da forma como a pergunta for feita, usando “Israel” e “Hamas” como palavras chaves mais frequentes, é muito possível que a ferramenta ofereça, ao aplicar o algoritmo sem qualquer

juízo de valor ou bom senso, apenas pela quantidade de vezes em que tais palavras chaves aparecem na Internet, uma resposta pela qual se conclua que os argumentos e estratégias de Israel para iniciar e manter a ofensiva são razoáveis., e vice-versa

Da mesma forma, isso poderia ocorrer com respostas que tragam referências a conteúdos homofóbicos, antissemitas, nazistas, racistas, promover fake news e quaisquer outros conteúdos ilegais, antiéticos, imorais, absurdos que podem muito bem levar um usuário mais desavisado a comprar tais ideias.

Neste caso, de quem seria a responsabilidade por tais eventos? Entendo que tanto da empresa cliente que está fazendo uso da ferramenta (que se furtou a adotar protocolos que permitissem uma conferência ou dupla validação para respostas envolvendo temas sensíveis) quanto do fabricante da ferramenta de IA (por oferecer uma ferramenta ainda até certo ponto insegura e imprecisa), em responsabilidade solidária ou mesmo subsidiária.

Diante deste tipo de situações e ao longo do tempo de uso de tais ferramentas, algumas fabricantes de ferramentas de IA buscaram otimizar a sua performance e procurar oferecer retornos e repostas mais seguras e confiáveis.

Neste sentido, muitas contrataram equipes moderadoras, ainda assim compostas por funcionários humanos, responsáveis por excluir da base de pesquisa destas ferramentas conteúdos considerados impróprios, ilegais, antiéticos, imorais, imprecisos, inverídicos ou até mesmo tendenciosos. Este é o caso, por exemplo, do ChatGPT, cuja desenvolvedora contratou grande equipe formada por trabalhadores quenianos e indianos, para tal tarefa, demonstrando nitidamente a ainda dependência do fator humano para uma melhor performance das ferramentas de IA.

Com o passar do tempo, hoje em dia, as empresas desenvolvedoras de ferramentas de IA perceberam que o ideal é trabalhar numa melhor e mais consciente programação dos algoritimos de machine learning das ferramentas, fazendo com que elas mesmas consigam, automaticamente, excluírem de suas bases conteúdos indesejados, através de melhores mecanismos de interpretação de texto e imagens, e complementando os conceitos de machine learning com quais tipos de conteúdo a ferramenta não deve lidar ou deve ignorar ou ainda não se tornar tendenciosa. Hoje em dia já existem exemplos de plataformas capazes de performar tais tarefas, como a Azure, da Microsoft (veja funcionalidades em Content Safety Studio – Microsoft Azure).

Mesmo assim e independentemente de tais plataformas tenham sofrido melhorias e otimização de performance de machine learning, sempre é extremamente importante que as

áreas de pesquisa e desenvolvimento tenham a total consciência da necessidade de criação e oferta de produtos que operem nos mais altos níveis de ética, moral, legalidade, confiabilidade e responsabilidade, ainda que não exista (ao menos neste momento, no Brasil) uma regulamentação especifica sobre IA. Isto pode ser operacionalizado, por exemplo, por treinamentos multidisciplinares (sempre envolvendo as áreas Legais, Privacidade e Proteção de Dados e CiberSegurança) aos profissionais de P&D e demais usuários deste tipo de ferramentas, determinados padrões de testes internos das ferramentas antes de seu uso efetivo, a serem levados a cabo pelos cientistas de dados responsáveis ou cursos técnicos específicos.

Sobre a IA Responsável, existe uma interessante matéria da Revista Época Negócios, que traduz bastante este conceito, de forma clara e didática: O que é IA responsável? – Época Negócios | Colunas (globo.com).

Contudo, apesar disso e mesmo se levando em consideração princípios de ética, responsabilidade, legalidade e moralidade aplicados à IA, neste momento, início de 2024, por mais desenvolvidos que estejam os algoritmos e protocolos de machine learning, estes ainda são responsivos, ou seja, somente conseguem trabalhar com situações já pré-programadas, que já aconteceram antes ou foram imaginadas previamente por seus desenvolvedores, sem a capacidade de reagir, posicionar-se ou criar respostas relacionadas a situações imprevistas, por não possuírem (ao menos não agora) a capacidade da interpretação analítica aliada à criatividade, à ligação de um repertório de possibilidades que permita à ferramenta reagir a situações inusitadas e imprevistas. Digamos que qualidades como interpretação analítica, criatividade, repertório e bagagem de vida são características diretamente relacionadas, hoje em dia, ao comportamento humano, com seu "jogo de cintura".

Quiçá, no futuro, a tecnologia evolua e consiga incluir estas características como parte dos requisitos dos algoritmos e do machine learning. Aí sim, passaremos a ter ferramentas de IA totalmente confiáveis, completas e com capacidade de operar independentemente. Neste momento, é possível que determinadas profissões sejam "ameaçadas" de maneira mais frontal mas é complexo cravar algo de maneira tão enfática. Tema certamente a ser explorado mais adiante em novo artigo.

Por enquanto, não se nega de maneira nenhuma a facilidade e maravilhas que a ferramenta pode trazer aos presentes dias. Contudo, colocá-la em operação com total e indistinta confiança para qualquer situação, sem que o usuário efetue uma checagem minuciosa do que foi oferecido como resposta me parece muitíssimo arriscado. A ferramenta deveria sim ser usada como suporte e otimização e aceleração de processos mais corriqueiros,

automáticos e maçantes, mas sempre sob revisão de seres humanos preparados para frear, antecipadamente, quaisquer respostas indesejadas, imorais, antiéticas ou ilegais que ela possa oferecer ou mesmo ter certeza de que a resposta oferecida não é mera cópia de trabalho de terceiro anteriormente produzido, ferindo seus direitos autorais sem mesmo que o consulente o perceba.