DOI 10.47402/ed.ep.c202320952201

CAPÍTULO 2

INTELIGÊNCIA ARTIFICIAL UTILIZADA PARA GARANTIA DE DIREITOS

João Pedro Albino
Docente e pesquisador do Programa de Pós-Graduação em Mídia e Tecnologia (Doutorado) na Universidade Estadual Paulista "Júlio de Mesquita Filho", Faculdade de Arquitetura, Artes, Comunicação e Design - Câmpus de Bauru.
Ana Cláudia Pires Ferreira de Lima
Discente do Programa de Pós-Graduação em Mídia e Tecnologia (Doutorado) na Universidade Estadual Paulista "Júlio de Mesquita Filho", Faculdade de Arquitetura, Artes, Comunicação e Design - Câmpus de Bauru.

RESUMO
A inteligência artificial está sendo usada na extração de dados digitais para comprovação de fatos em processos judiciais para garantia de direitos, como os direitos fundamentais à honra e a não discriminação e os direitos sociais ao reconhecimento de uma relação de emprego e ao pagamento de horas extras, dentre outros. Este artigo destaca a utilização da Inteligência Artificial na produção de provas digitais em diferentes contextos, como a identificação de autores de postagens nas mídias sociais, análise de metadados de fotos digitais, a realização de pesquisa patrimonial de devedores em processos de execução e a análise de mensagens de texto para identificar informações relevantes em investigações criminais. Dentre as vantagens do uso da Inteligência Artificial na análise de dados, tem-se a capacidade de processar grandes quantidades de dados rapidamente e a identificação de padrões e tendências que podem não ser facilmente detectáveis por humanos. A volatilidade dos dados digitais impõe ao profissional do Direito conhecimentos gerais de computação forense e algumas ferramentas de inteligência artificial para coletar e armazenar esses dados para assegurar a validade da prova em processo judicial. Neste artigo serão apresentadas algumas ferramentas de inteligência artificial utilizadas para análise de dados forenses, que auxiliam na extração das provas digitais. Para elaboração deste artigo foi realizada pesquisa em artigos científicos, livros e materiais de cursos e palestras.
PALAVRAS-CHAVE: ciência de dados; mídia e tecnologia; inteligência artificial; provas digitais; computação forense.

1 INTRODUÇÃO

O crescimento do número de usuários da internet nos últimos anos gerou um aumento exponencial do número de dados registrados no espaço cibernético. Relatório do Digital Global Overview Report revelou a existência de 5,16 bilhões de usuários da internet em janeiro de 2023, representando 64,4 % da população mundial. Destaca, ainda, que a média de uso diário da internet no mundo é de 6h37 e no Brasil é de 9h32[1]. O ser humano não consegue processar

[1] Digital 2023 July Global Statshot Report (Janeiro 2023). Publicado em 20/02/23. Disponível em: https://www.slideshare.net/DataReportal/digital-2023-global-overview-report-summary-version-january-2023-v02 Acesso em: 27/03/23.

esse grande volume de dados sozinho, precisando do auxílio da máquina, através da inteligência artificial.

Caselli (2022, p. 41) nos revela os seguintes dados:

> Para se ter ideia da quantidade do volume de informações produzidas em apenas um dia na internet e, especificamente nas redes sociais, o site influencer marketing hub, especializado em social analytics aponta que a cada 24 horas, 7 bilhões de vídeos são visualizados na plataforma do Youtube, 729 milhões de tweets são realizados. Na rede socias Instagram, 78 milhões de fotos são carregadas, 432 mil stories são postados e 2 bilhões de curtidas sao realizadas. Já na rede social Facebook, 4 bilhões de publicações são realizadas. Na plataforma do Google, 6 bilhões de pesquisas são realizadas e 27 bilhões de mensagens são enviadas via WhatsApp.

O grande volume de dados explica a razão de a reprodução de mensagens de WhatsApp ser uma das provas muito comum apresentadas nos processos que tramitam na Justiça do trabalho. Justamente porque é o meio de comunicação digital mais utilizado. Mas a prova digital deve ser reproduzida corretamente perante o Poder Judiciário, sob pena de não ter validade.

Os dados digitais armazenados podem ser utilizados como prova em processos judiciais, desde que observadas as regras do devido processo legal. “O processo de identificação, preservação, coleta e apresentação de evidência digital que possa ser apresentada em processo judicial é conhecido como computação forense.” (RAJASEKAR et al., 2023).

O crescimento exponencial dos dados digitais impõe o uso de ferramentas de inteligência artificial para auxiliar na coleta e análise desses dados, para identificação e apresentação de provas digitais.

Morais et al. (2018, p. 13-14) assim conceituam Big Data:

> A princípio, podemos definir o conceito de Big Data como conjuntos de dados extremamente amplos e que, por esse motivo, necessitam de ferramentas preparadas para lidar com grandes volumes de dados, de forma que toda e qualquer informação nesses meios possa ser encontrada, analisada e aproveitada em tempo hábil. Com o aumento significativo da quantidade de dados gerados pela internet com o surgimento das mídias sociais, é necessário gerenciar e armazenar as informações de maneira organizada. Esses dados podem ser classificados em estruturados, não estruturados e semiestruturados com base no seu gerenciamento e armazenamento.

Antunes e Rodrigues (2018, p. 30) explicam que o termo *Big Data*:

> refere-se ao conjunto gigantesco de dados que podem ser recolhidos e analisados computacionalmente, com o objetivo de identificar padrões, associações e tendências relacionadas com um determinado negócio ou atividade.

O fenômeno do *Big Data* e o avanço da tecnologia permitiu o avanço do uso da inteligência artificial para identificar determinados padrões e para automatização de tarefas que seriam difíceis de serem realizadas por humanos. *Data mining* utiliza técnicas de *machine*

learnig para extrair conhecimento dos dados para tomada de decisões (ESCOVEDO e KOSHIYAMA, 2020, p.13).

A inteligência artificial está sendo cada vez mais utilizada na produção de provas judiciais, transformando a forma com advogados e juízes lidam com a análise do conhecimento extraído dos dados digitais e a tomada de decisões.

Belluzzo e Valente (2021, p.23) assim definem a Inteligência Artificial:

> é a área de pesquisa da computação dedicada a buscar métodos ou dispositivos computacionais que possuam as características básicas: capacidade de raciocínio (aplicar regras lógicas a um conjunto de dados disponíveis para chegar a uma conclusão), aprendizagem (aprender com os erros e acertos de forma a no futuro agir de maneira mais eficaz), reconhecer padrões (tanto padrões visuais e sensoriais, como também padrões de comportamento) e inferência (capacidade de conseguir aplicar o raciocínio nas situações do nosso cotidiano).

Escovedo e Koshiyama (2020, p. 13) conceituam Ciência de dados, que utiliza inteligência artificial para análise dos dados:

> Refere-se à coleta de dados de várias fontes para fins de análise, com o objetivo de apoiar a tomada de decisões, utilizando geralmente grandes quantidades de dados, de forma sistematizada. Quase sempre, além de olhar para os dados passados para entender o comportamento dos mesmos (atividade conhecida como Business *Intelligence* - BI), deseja-se também realizar análises de forma preditiva, por exemplo, utilizando técnicas de Data Mining e/ou *Machine Learning*.

Stanton (2012, p. 4 e 11) nos ensina que a Ciência de Dados "refere-se a uma área emergente de trabalho preocupada com a coleta, preparação, análise, visualização, gerenciamento e preservação de grandes conjuntos de informações." E nos ensina que o maior objetivo do cientista de dado é ajudar as pessoas a transformar dados em informação, conhecimento, entendimento e sabedoria, referindo-se à Tim Berners-Lee, o inventor da internet, ou seja, é extrair o máximo possível dos dados, que seja útil para determinado campo de aplicação.

No âmbito jurídico, ferramentas de aprendizado de máquina podem ser utilizadas para analisar grandes volumes de dados de dispositivos eletrônicos, como computadores e telefones celulares, para análise de dados forenses, em busca de provas relevantes para comprovação de um fato em uma ação judicial, na qual se pleiteia o reconhecimento de um direito. Isso permite que os advogados encontrem informações valiosas de forma mais rápida e precisa do que seria possível manualmente.

Existem diversas ferramentas disponíveis para extração de provas digitais. Algumas com versões gratuitas e pagas. A inteligência artificial, com técnicas de aprendizado de máquina, tem contribuído para a automação dessas ferramentas. Neste artigo serão apresentadas

algumas ferramentas de inteligência artificial utilizadas para análise de dados forenses, que auxiliam na extração das provas digitais.

2 CIÊNCIA DE DADOS E COMPUTAÇÃO FORENSE

De nada adianta o aprendizado de máquina, se ele não for aplicado de forma prática na busca de soluções para os problemas humanos. Assim, o trabalho do Cientista de Dados na organização, coleta, análise e arquivamento é essencial para o resultado útil do aprendizado de máquinas.

Até 2010 o desafio era armazenar os dados, o que foi resolvido com novas tecnologias, a exemplo das plataformas de software de computação distribuídas, como a *Hadoop*. Atualmente, o foco voltou-se para o processamento dos dados.

Os dados podem ser classificados como estruturados, não-estruturados e semiestruturados, conforme explanado por Sharma (2023):

> **Dados estruturados** são aqueles que estão organizados em uma estrutura rígida, a qual foi previamente planejada para armazená-los. Geralmente os dados estruturados estão organizados em forma de colunas e linhas, como em uma tabela ou planilha eletrônica, por exemplo.
> Já os **dados não-estruturados** são aqueles que são armazenados utilizando-se uma estrutura flexível e dinâmica ou sem uma organização definida. O exemplo mais comum de dado não estruturado é um documento ou arquivo contendo imagens (gráficos e fotos) misturado com textos.
> As *redes sociais*, que apresentam elevado volume de dados criados diariamente pelos usuários, representam outro exemplo de **dados não estruturados**.
> Atualmente, mais de 80% do conteúdo digital gerado no mundo é do tipo **não estruturado.**
> Os **dados semiestruturados** apresentam uma representação *heterogênea*, ou seja, possuem estrutura, mas esta é flexível. Esta representação de dados agrega um pouco dos formatos estruturado e não-estruturado em termos de benefícios. Facilita o controle por ter um pouco de estrutura, mas também permite uma maior flexibilidade.

Sharma (2023) previu que a partir de 2020 “a maioria dos dados disponíveis são não-estruturados. Esses dados são gerados de diferentes fontes, como registros financeiros, arquivos de texto, formulários multimídia, sensores e instrumentos.”

Ferramentas de inteligência de negócios não bastam para processar esse grande volume e variedade de dados, havendo necessidade de “ferramentas e algoritmos analíticos mais complexos e inovadores para poder processar, analisar e extrair conhecimentos (*insights*) relevantes.”

Sharma (2023) também distingue o *Business Intelligence* (BI) do Data Science (Ciência de Dados):

> O BI analisa basicamente dados passados (dados históricos) para compreender uma situação ou determinado evento *após a sua ocorrência* de forma a obter *insights* e

> apresentar tendências nos negócios. O BI permite coletar dados de fontes externas e internas, prepará-los, realizar consultas e criar **dashboards6** para responder a perguntas como "como está a análise de receita trimestral" ou "mostrar quais problemas de curto e/ou longo prazo nos negócios". O BI pode avaliar o impacto de determinados eventos em curto prazo.
> A Ciência de Dados é uma abordagem mais prospectiva, de forma exploratória, com foco na *análise de dados passados* ou *atuais* e *na previsão de resultados futuros* com o objetivo de tomar decisões com conhecimento. A Ciência de Dados busca responder *perguntas abertas* como "**quais**" e "**de que forma**" os eventos ocorrem.

A Ciência de dados é a combinação de ferramentas, algoritmos e princípios de aprendizado de máquina (ML) com o objetivo de *descobrir padrões ocultos* a partir de *dados brutos* (SHARMA, 2023, p. 4).

A ciência de dados está sendo aplicada em vários setores, a exemplo das empresas para entender a necessidade de seus clientes, através de seus históricos de navegação, compras, idade e renda, utilizando o algoritmo para traçar o perfil do cliente e verificar os produtos que podem ser mais propícios para oferecer, de acordo com seus interesses, tornando os negócios mais rentáveis.

A análise de dados também é muito utilizada na área jurídica, em diversas disciplinas, a exemplo da jurimetria e da computação forense.

Nunes (2020) define Jurimetria:

> como a disciplina do conhecimento que utiliza a metodologia estatística para investigar o funcionamento de uma ordem jurídica. A partir dela, fica claro que a Jurimetria se distingue das demais disciplinas jurídicas tanto pelo objeto como pela metodologia empregada na sua análise.

A computação forense, que também requer a análise de dados digitais, *"consiste no uso de métodos científicos na preservação, coleta, validação, identificação, análise, interpretação, documentação e apresentação de evidências digitais"*, conforme definição de Pinheiro (2021, p. 280).

Antunes e Rodrigues (2018, p. 135) assim conceituam a computação forense:

> A disciplina de **computação forense**, também vulgarmente designada por informática forense ou, ainda, forense digital, é uma das mais recentes disciplinas de ciência forense. Consiste essencialmente na aplicação de técnicas científicas e boas práticas de manuseamento de equipamentos informáticos, com vista à obtenção de **provas digitais**, ou **evidências**, com **validade probatória** perante um tribunal.

3 INTELIGÊNCIA ARTIFICIAL

Inteligência artificial é um ramo da ciência da computação que se concentra em criar máquinas que podem executar tarefas humanas. A Inteligência artificial abrange o estudo de *Machine Learning* (aprendizado de máquina) e também de *Deep Learning* (aprendizagem

profunda). *Machine learning* é o estudo científico de algoritmos e modelos estatísticos para executar uma tarefa usando inferência em vez de instruções.[2]

O *Machine learning* ou aprendizado de máquina se concentra no uso de dados para treinar modelos de aprendizado de máquina, para que esses modelos possam fazer previsões. Russel e Norvig (2022, p. 1.544) assim explicam:

> Quando o agente é um computador, nós o chamamos aprendizado de máquina: um computador observa alguns dados, monta um modelo baseado nos dados e usa o modelo como uma hipótese sobre o mundo e um software que pode resolver problemas.

Como exemplos da utilização do aprendizado de máquina temos: a) o filtro de Spam, que é um programa de aprendizado de máquina que foi treinado com exemplos de spam e mensagens de e-mail regulares; b) detecção de compras fraudulentas com cartão de crédito, detectadas pelo programa de aprendizado de máquina, treinado com exemplos de compras fraudulentas e de transações legítimas; c) recomendações de livros, filmes, músicas e produtos, com a utilização do treinamento do programa de aprendizado de máquinas com dados de hábitos e compras do usuário e de outros consumidores; d) identificação de tumores em exames médicos e e) agrupamento de fotos em mídias sociais através do reconhecimento facial.

Martins (2022, p. 505) ressalta o uso da IA para automação de tarefas repetitivas na área jurídica:

> As bases foram sendo lançadas durante a segunda metade do século passado, mas foi o início do segundo milénio que colocou os juristas perante um já consideravelmente extenso acervo de ferramentas informáticas que, utilizando IA, desanuviam os juristas de morosas e repetitivas tarefas. Em especial, as técnicas de processamento de linguagem natural associadas a estatística e probabilidades são tremendamente úteis e promissoras.
> O *software* de *legal analytics* permite analisar informação contida em documentos selecionando a parte relevante, ou seja, a que merece a atenção do jurista. A sua aplicabilidade em *due diligences*, na preparação de um julgamento ou na investigação criminal está à vista.
> A revolução sobre a ponderação do risco do litigio é protagonizada pelos programas de justiça preditiva. O utilizador introduz no sistema os dados essenciais sobre a ação que pretende propor e fica a saber qual é a probabilidade de obter vencimento, e em que medida (se o pedido for quantificado). Um obstáculo técnico à sua disseminação reside na necessidade de disponibilidade dados. Em países ou contextos (arbitragem) em que a informação judicial não seja livremente acessível, mais difícil (embora não inviável) se torna implementar estes sistemas.

Há três tipos principais de aprendizado de máquina: aprendizagem supervisionada, aprendizagem não supervisionada e aprendizagem por reforço.

[2] AWS Academy Machine Learning Foundations. Module 02 Student Guide, p. 11.

Na aprendizagem supervisionada, um modelo usa entradas e saídas conhecidas para generalizar saídas futuras, conforme explicado por Russel e Norvig (2022, p. 1.548):

> Por exemplo, as entradas poderiam ser imagens de câmera, cada uma acompanhada por uma saída dizendo "ônibus" ou "pedestre" etc. Uma saída como essa é chamada rótulo. O agente aprende uma função que, quando recebe uma nova imagem, prevê o rótulo apropriado.

Na aprendizagem não supervisionada, o modelo não conhece as entradas ou as saídas. Ele encontra padrões nos dados sem auxílio.

> No aprendizado não supervisionado, o agente aprende padrões na entrada, embora não seja fornecido nenhum feedback explícito. A tarefa mais comum de aprendizagem não supervisionada é o agrupamento: a detecção de grupos de exemplos de entrada potencialmente úteis. Por exemplo, quando recebe milhões de imagens tomadas da Internet, um sistema de visão computadorizado pode identificar um grande grupo de imagens semelhantes que alguém chamaria de "gatos" (RUSSEL e NORVIG, 2022, p. 1.549).

O terceiro tipo é a aprendizagem por reforço, em que o modelo interage com seu ambiente e aprende a tomar ações que maximizam as recompensas.

> No aprendizado por reforço, o agente aprende a partir de uma série de reforços: recompensas e punições. Por exemplo, no fim de um jogo de xadrez o agente é informado de que ele ganhou (uma recompensa) ou perdeu (uma punição). Cabe ao agente decidir quais das ações anteriores ao reforço foram as maiores responsáveis por isso e alterar suas ações visando a mais recompensas no futuro (RUSSEL e NORVIG, 2022, p. 1.549).

Saxena et al. (2023, p. 135 e 138) falam sobre a utilização de *Machine Learning* (Aprendizado de Máquina) e *Deep Learning* (Aprendizado Profundo) nas investigações forenses :

> O aprendizado de máquina (AM) tem sido frequentemente usado em investigações forenses digitais para descoberta de dados, triagem de dispositivos e forense de rede. Especificação de tarefas, criação de recursos e avaliação e otimização são as três fases em aplicativos de AM. Dependendo do tipo de rótulos de destino, uma tarefa de AM pode ser classificada como uma classificação/agrupamento ou um desafio de regressão. As transformações e seleções de recursos são feitas durante o experimento para reduzir o over-fitting, melhorar o desempenho e reduzir o tempo de treinamento. Da mesma forma que os humanos não desenvolvem recursos, as aplicações forenses digitais de aprendizado profundo (AP) são análogas ao aprendizado de máquina (AM). Em vez disso, uma técnica de aprendizado de uso geral é usada para aprender com os dados. Otimização e inferência são as duas etapas de um modelo AP. Os valores que conectam as unidades denominadas neurônios especificados no modelo são alterados por meio da fase de treinamento. A interpretação é usada para fazer previsões precisas com base em dados não rotulados, não visíveis durante o aprendizado. (....) O aprendizado profundo é um ramo da pesquisa de aprendizado de máquina que resolve problemas analisando grandes conjuntos de dados e empregando redes neurais. De acordo com Jackson (2019), uma rede neural é uma representação muito simplificada de uma rede cerebral biológica, retratada como uma coleção de "neurônios artificiais" interconectados. Essas redes fazem julgamentos com base na entrada de dados e modificam valores dependendo do feedback para se aproximar da saída pretendida.
>
> Devido à sua capacidade de adaptação sem exigir entrada do usuário, a inteligência artificial é benéfica em várias situações. Muitas soluções recomendam o uso de

> aprendizado de máquina para lidar com os enormes volumes de dados que a perícia digital exige.

É importante conhecer os diferentes tipos de Aprendizado de Máquina, os quais orientam na seleção de algoritmos que fazem sentido para resolver o problema de negócio.

4 FERRAMENTAS DE INTELIGÊNCIA ARTIFICIAL APLICADAS AO DIREITO

Este artigo visa destacar a utilização de ferramentas de inteligência artificial aplicada para garantia de direitos, a exemplo de produção de provas em processos judiciais. Destacaremos algumas ferramentas de inteligência artificial que auxiliam na análise dos dados forenses, com uma breve explanação de sua finalidade, sem a intenção de esgotar o tema.

Mister destacar que fontes de dados e de informação podem ser classificadas em fontes abertas, de livre acesso aos usuários da internet e fontes fechadas, cujo acesso depende de autorização prévia, a exemplo de login e senha ou ordem judicial.

Caselli (2022, p. 34 e 39) assim conceitua as fontes abertas:

> Compreendemos, então, que fonte aberta é todo o meio de busca de informações que estejam livremente dispostas, ou seja, que não estejam em bases protegidas, que demandem senhas para seu acesso, intervenção judicial ou manobra técnica. Por exemplo, seria a informação disposta em um site, acessível a qualquer internauta que busque por aquele conteúdo. [...] Entendemos que fonte aberta é todo dado, informação ou conhecimento livremente disponibilizado por seu titular ou de quem lhe faça as vezes, atribuindo-lhes, assim o caráter de publicidade voluntariamente, e que são capazes de produzir conhecimento ou prova em procedimento administrativo ou judicial.

Fontes fechadas, a contrário senso, são aquelas cujo acesso depende previamente de autorização, a exemplo de login e senha ou ordem judicial. Como exemplo de fontes fechadas que requerem login e senha para acesso temos os dados de contas bancárias, e-mail, cadastros particulares, prontuários médicos, perfis das mídias sociais que não estejam no modo público, ou seja, requerem autorização prévia de acesso de seu titular.

Alguns exemplos de fontes fechadas que dependem de determinação judicial para serem acessadas são as declarações de imposto de renda e outras informações cadastrais de pessoas físicas e jurídicas requisitadas pelo Poder Judiciário à Receita Federal, com o afastamento do sigilo fiscal. Dados de Geolocalização captados pelas Operadoras de telefonia celular também são fontes fechadas.

Vários fatos que se tornam relevantes para o mundo jurídico podem ser comprovados através dos dados digitais, obtidos de fontes abertas ou fechadas. Esse conhecimento é essencial aos advogados, que devem verificar se há registros digitais dos fatos jurídicos que se pretende provar e qual a melhor forma de comprová-los nos autos do processo judicial. Caso os dados

que se pretende utilizar como prova sejam provenientes de fontes fechadas, o advogado deve requerer o respectivo acesso ao seu titular ou ao Judiciário, para que sua prova não seja considerada ilícita, ou seja, obtida por meios não permitidos pela legislação.

Diversos fatos que podem ser relevantes para o mundo jurídico podem ser registrados em dispositivos de armazenamento físicos como pen drives, discos rígidos, drives SSD, dispositivos móveis ou discos de armazenamento ópticos ou em área de armazenamento na nuvem.

As relações humanas ocorrem cada vez mais no meio digital, onde os fatos são registrados. Há muitos casos de crimes contra a honra (injúria, difamação e calúnia) perpetrados nas mídias digitais, a exemplo de *facebook*, *instagram* e *twitter*, sendo necessário verificar quem de fato postou a(s) mensagem (s) ofensiva (s), para aplicação da sanção correspondente, a pedido da vítima, tanto na esfera criminal como na esfera cível (indenização por danos morais). Para extração desses dados há necessidade de conhecimento de direito digital e de direito probatório digital, com algumas técnicas de computação forense.

Pela análise de registros digitais também é possível apurar a responsabilidade de vazamento de dados de uma empresa, a exemplo de constatação de quem realizou a cópia dos dados e de quem os enviou por e-mail, ou, até mesmo, se houve indução da pessoa em erro através do *phishing*, que é uma forma de engenharia social da qual uma pessoa má intencionada se utiliza para ganhar a confiança da vítima e insere um link falso para captação dos dados.

Ferramentas de inteligência artificial permitem a extração e análise de dados forenses de forma mais rápida e precisa do que seria possível manualmente, auxiliando na produção de provas digitais em processos judiciais.

4.1 Detecção de Metadados de Documentos

Metadados são dados sobre dados, *"são informações estruturadas que auxiliam na descrição, identificação, gerenciamento, localização, compreensão e preservação de documentos digitais, além de facilitar a interoperabilidade de repositórios"*[3]

[3] ONLINE COMPUTER LIBRARY CENTER (2002). Preservation metadata and the OAIS information model: a metadata framework to support the preservation of digital objects (PDF). Ohio, USA: OCLC. 51 páginas. Acesso em: 28 de novembro de 2019, *apud* Tribunal Regional do Trabalho da 7ª Região. MSCiv 0080297-42.2021.5.07.0000. Desembargadora Maria Roseli Mendes Alencar. 10/06/2021. Disponível em: file:///C:/Users/Global/Downloads/Documento_5b9bc6e.pdf. Acesso em: 27/03/23.
The OCLC/RLG Working Group on Preservation Metadata. Preservation Metadata and the OAIS Information Model. A Metadata Framework to Support the Preservation of Digital Objects. Junho/2002. Disponível em: https://www.oclc.org/content/dam/research/activities/pmwg/pm_framework.pdf Acesso em: 27/03/23.

Metadados são dados que fornecem informações sobre outros dados. Eles descrevem as características e propriedades dos dados, como formato, autor, data de criação, localização, entre outros atributos. Em outras palavras, são informações que fornecem contexto e ajudam a entender e gerenciar melhor os dados. Os metadados são usados em diversas áreas, incluindo em sistemas de gerenciamento de conteúdo, bibliotecas digitais, arquivos de mídia e em aplicações de inteligência artificial, onde são usados para treinar e aprimorar os modelos de aprendizado de máquina.

Todos os arquivos digitais possuem metadados. Através deles é possível verificar a origem, data, IP e geolocalização do local em que o documento foi produzido. É possível inserir dados em documentos digitais para preservação dos direitos autorais de seu autor.

Ferramentas de IA podem ser usadas para extrair e analisar os metadados, como informações de data e hora, localização e dispositivo, para produzir provas digitais. Isso pode incluir análise de metadados de imagens e de arquivos para investigar a origem, integridade e autenticidade do documento.

Dentre os softwares utilizados para analisar dados, Oetinger (2021, p. 56-57) destaca a diferença entre os *softwares* comerciais e de ferramentas *open source*:

> Os fornecedores deixam os *softwares open source* disponíveis gratuitamente a qualquer um que os utilize. Em geral, não há restrições quanto ao seu uso; você pode utilizá-los com propósitos educacionais, financeiros ou para testes. O aspecto positivo é que o software estará disponível sem custo algum na maioria das situações. A desvantagem é que você terá pouco ou nenhum suporte técnico caso algo dê errado. Sua opção dependerá totalmente de seu conjunto de habilidades e de seu nível de conforto para trabalhar com essas ferramentas. Muitas ferramentas *open sour*ce utilizam uma CLI (*Command-Line Interface*, ou Interface de Linha de Comando) em vez de ter uma GUI (*Graphical User Interface*, ou Interface Gráfica de Usuário), e isso pode intimidar os novos usuários.
> Em geral, uma ferramenta comercial terá um melhor serviço de atendimento ao cliente, documentação e atualizações periódicas. A desvantagem é que você pagará por esses serviços. Na verdade, para qualquer tarefa que uma ferramenta forense comercial possa fazer, há uma ferramenta *open source* capaz de fazer o mesmo. Uma ferramenta comercial será capaz de desempenhar diversas funções, ao passo que, com um framework *open source*, talvez você tenha de usar uma ou mais ferramentas distintas para executar a mesma tarefa

Há vários meios de extração de metadados dos documentos digitais, a exemplos de fotografias digitais, arquivos de textos, e-mails, postagens em mídias sociais, áudio e vídeo. Os metadados dos arquivos digitais devem ser apresentados nos autos do processo para comprovação da autenticidade e integridade do documento.

Ao fazer uma postagem nas mídias sociais, vários metadados são compartilhados com o Provedor de Aplicação, a exemplo do endereço IP, data e horário da postagem, dados da localização de onde está sendo feita a postagem.

A Figura 1 mostra a quantidade de dados que compartilhamos através de e-mails, pesquisa em motores de busca da internet, telefonemas, fotos digitais e uso das mídias sociais, devendo tudo estar descrito nos termos de uso e na política de privacidade de dados dos provedores de aplicação, os quais nem sempre são lidos por seus usuários, que realizam um contrato de adesão às suas cláusulas ao utilizarem o serviço:

Figura 1– Dados compartilhados por e-mail e mídias digitais.

Fonte: *The Guardian* (2013).

4.2 Análise de Áudio

O reconhecimento de voz e de imagem também está sendo utilizado cada vez mais no direito. Ferramentas de inteligência artificial podem ser usadas para identificar indivíduos em gravações de áudio e vídeo, o que é útil em investigações criminais e processos civis.

Diante da impugnação a trecho de áudio apresentado como prova em processo judicial, há a necessidade de realização de perícia judicial para verificação de sua autenticidade.

Ferramentas de *IA* podem ser usadas para analisar áudios, como gravações de telefonemas e aplicativos de mensagens instantâneas para verificação da autenticidade e integridade do áudio. Isso pode incluir transcrição automática (para facilitar a análise do conteúdo), reconhecimento de fala (através de perícia judicial, que pode detectar a

autenticidade, ou seja, a vinculação da fala ao seu verdadeiro autor através de sua frequência) e análise de sentimento para determinar o tom da conversa, verificando-se o contexto.

Dentre os softwares utilizados para análise da autenticidade de áudios há o *Audacity* e o *Sound Forge*.

4.3 Análise de Vídeo

Ferramentas de IA podem ser usadas para analisar vídeos, como gravações de câmeras de vigilância, para extrair informações relevantes, a exemplo da data e local em que ocorreu determinado fato e os atores nele envolvidos. Isso pode incluir reconhecimento facial, reconhecimento de placas de veículos e rastreamento de objetos em movimento.

4.4 Análise de Texto

Ferramentas de IA podem ser usadas para analisar textos, como mensagens de texto e e-mails, para extrair informações relevantes. Isso pode incluir análise de sentimento, detecção de linguagem inapropriada e extração de informações pessoais.

A análise de linguagem natural, que permite que os computadores entendam e processam textos escritos em linguagem natural, também é uma técnica cada vez mais utilizada no direito para analisar conversas e mensagens de texto. Além disso, sistemas de recomendação de jurisprudência estão sendo desenvolvidos para ajudar advogados e juízes a encontrar decisões relevantes para seus casos.

O cabeçalho de um e-mail apresenta diversos metadados, a exemplo do ID da mensagem, data e hora de envio do e-mail e da sua entrega, dentre outros dados, conforme detalhado pelo site King Host[4]:

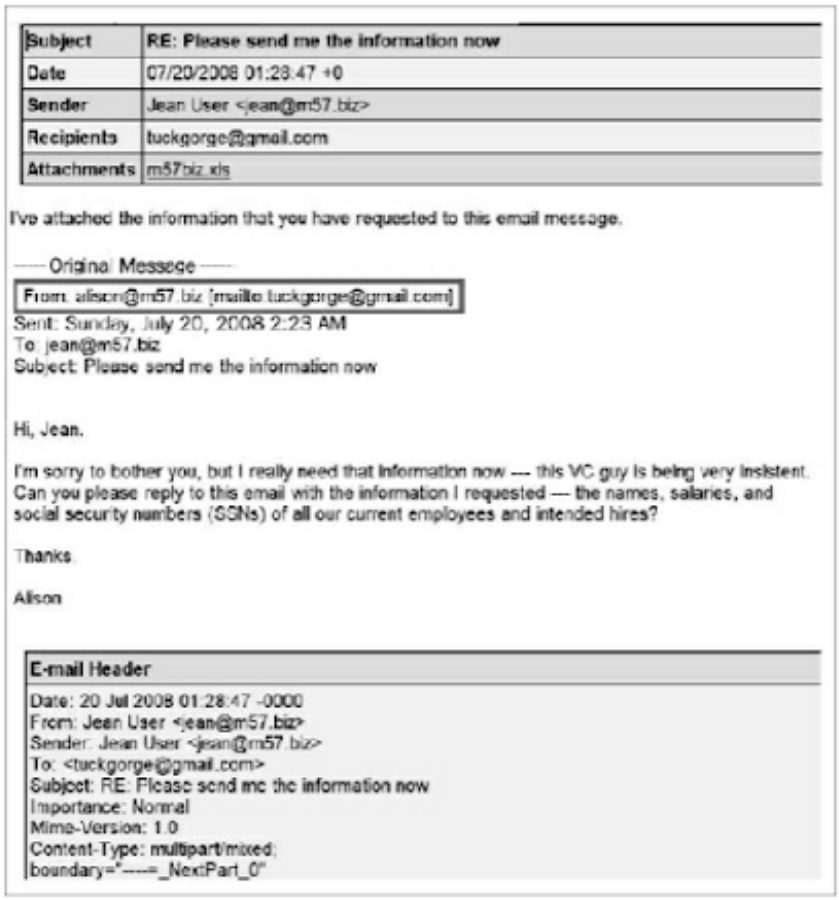
> O cabeçalho de uma mensagem, também conhecido como header, é um registro de informações de um e-mail. Através dele podemos consultar diversas informações pertinentes à mensagem, como: data e hora de envio, data e hora de entrega, servidor SMTP que enviou a mensagem, todos os servidores por onde a mensagem trafegou, servidor de recebimento da mensagem, *from* (quem envia a mensagem), *to* (quem deve receber a mensagem), filtros de spam e outras informações não menos importantes.

Em um estudo de caso, Jean, colaborador de uma empresa, era suspeito de ter enviado dados sigilosos empresariais para terceiro. Jean explicou que enviou a planilha solicitada à sua superior, Allison. Oettinger (2021, p. 222-223) cita como esse vazamento de dados foi elucidado a partir da análise dos metadados de e-mails:

[4] Analisador de cabeçalho (header) de e-mail - kingHost. https://King.host/wiki/ferramenta-header/ Acesso em: 26/03/23.

Vemos que Jean enviou o email para o que parece ser allison@M57.biz, mas, na verdade, ele foi enviado para tuckgorge@gmail.com. Podemos, então, filtrar por tipo de arquivo – nesse caso, os arquivos .eml – e veremos o resultado a seguir (Figura 3??): Se observarmos as colunas Sender (Remetente) e Recipients (Destinatários) e colocarmos os dados em ordem cronológica, poderemos ter uma boa ideia da comunicação por email entre o invasor e Jean. Parece que houve um comprometimento da conta de Allison, pois podemos ver o nome "Alex" e a conta de email tuckgorge@gmail.com associada à conta dela.

Figura 2 – Análise de metadados de e-mail.



Fonte: Oettinger (2021, p. 223).

Oettinger (2021, p. 223-225) continua a explicação:

Utilizar o recurso de lista de eventos do *X-Ways Forensics* nos permite identificar exatamente quando o arquivo foi comprometido e a partir de qual vetor. Agora podemos direcionar nossa investigação para o computador de Allison e determinar se o invasor comprometeu seu sistema. Com base nesses resultados iniciais, acredito que Jean tenha sido vítima de um ataque de *phishing*.

Um aspecto de que eu gosto no *X-Ways Forensics* é a sua capacidade de reunir datas e horas de fontes de dados tradicionais e combiná-las com os artefatos em questão – nesse caso, os *emails*. Isso nos dá outro nível de especificidade e contexto para nossas investigações. [...] Tenho notado que as suítes forenses atualmente estão incluindo também uma análise de linha do tempo em seus produtos. Já discutimos o *X-Ways Forensics* e sua capacidade de criar uma linha do tempo para análise com a sua funcionalidade de lista de eventos. Incluí uma lista de algumas suítes forenses adicionais que você poderá usar para analisar dados de linhas do tempo. A lista a seguir não inclui todas as suítes forenses disponíveis:

• *Belkasoft Evidence Center: belkasoft.com/ec*

• *Autopsy: http://www.sleuthkit.org/autopsy*
• *Recon Lab: sumuri.com/software/recon-lab*
• *Paladin: sumuri.com/software/paladin*

O *X-Ways* não é a única ferramenta que pode ser usada para criar linhas do tempo; há também uma série de ferramentas *open source* que podem ser utilizadas. Uma das mais comuns é o *Plaso/log2timeline*, que será discutido a seguir.

4.5 Análise de Fotografias Digitais

Algoritmos de aprendizado de máquina podem ser usados para detectar imagens que foram manipuladas ou modificadas. Isso pode incluir técnicas como detecção de imagem reversa, que busca características únicas na imagem original para determinar se ela foi modificada.

A busca de imagens reversas na internet pode constatar o uso de imagens semelhantes postadas em outras mídias sociais, através das quais pode-se obter metadados essenciais para casos investigados. Por exemplo, uma pessoa, investigada de ter cometido um crime em determinado local, apresenta uma foto alegando que não estava naquele local investigado. Através de busca por imagem reversa os agentes policiais encontram foto semelhante postada em mídia social, podendo-se obter os metadados da foto, a exemplo da data e hora em que ela foi tirada e inclusive sua geolocalização.

Uma das ferramentas para extração de metadados das fotografias digitais é apresentada por Barreto, Wendt e Caselli (2017, p. 147):

> O exif metadata é a informação adicional do arquivo da fotografia que pode ter dados sobre data e hora, tamanho, características da câmera ou do smartphone, dados de luminosidade e outras informações úteis. Em alguns casos, quando o GPS (Global Positioning System) do equipamento está ligado, é possível obter a real posição em que a fotografia foi tirada. Cada metadado traz consigo dados individualizadores da imagem produzida.

Algumas ferramentas para extração de metadados de fotografias digitais podem ser encontradas nos seguintes sites: http://fotoforensics.com/ http://www.exif-viewer.com/ e http://www.pic2map.com/.

4.6 Identificação de Padrão

Algoritmos de aprendizado de máquina podem ser usados para identificar padrões e tendências em grandes conjuntos de dados, incluindo provas digitais. Isso pode incluir análise de dados financeiros para identificar fraudes, por exemplo em operações bancárias que fogem do padrão do perfil do cliente; análise de registros de rede para identificar atividades maliciosas, e análise de log de acesso para identificar comportamentos suspeitos.

Ferramentas de análise de dados podem ser usadas para analisar grandes conjuntos de dados, incluindo dados de redes sociais, dados de log de acesso, dados de transações financeiras e outros tipos de dados, para identificar tendências, padrões e relações entre diferentes tipos de dados.

4.7 Ferramentas de Cibersegurança

Essas ferramentas podem ser usadas para monitorar a atividade de rede, detectar ameaças e proteger contra-ataques cibernéticos. Elas podem ser usadas para investigar atividades maliciosas e identificar provas digitais relevantes.

4.8 Análise de Mídias Sociais

Ferramentas de IA podem ser usadas para analisar conteúdo de mídias sociais para produzir provas digitais. Isso pode incluir análise de sentimento para determinar a opinião das pessoas sobre um determinado assunto, análise de relações entre contas e usuários para identificar atividades maliciosas, e análise de conteúdo para identificar informações relevantes.

Essas são apenas algumas das ferramentas de inteligência artificial (IA) que podem ser utilizadas para produzir provas digitais em processos judiciais envolvendo mídias sociais. Cada ferramenta tem suas próprias capacidades e limitações, e é importante escolher a ferramenta certa para o caso específico. Além disso, é importante lembrar que o uso de IA deve ser feito de forma ética e justa, garantindo a privacidade e segurança dos dados e a transparência nas decisões tomadas pelos algoritmos:

4.8.1 *Social Bearing*

Social Bearing é uma ferramenta de inteligência artificial (IA) que permite analisar conteúdo de redes sociais para produzir provas digitais. Ele usa algoritmos de aprendizado de máquina para analisar dados de redes sociais, como postagens, comentários, curtidas e outras informações, para identificar tendências, opiniões, sentimentos e relações entre usuários. Ele pode ser usado para investigações criminais, monitoramento de atividades maliciosas, análise de risco e outras finalidades.

4.8.2 *Sowdust*

Sowdust é uma ferramenta de inteligência artificial (IA) que permite aos usuários automatizar tarefas de busca e análise de dados. Ele usa algoritmos de aprendizado de máquina para analisar grandes conjuntos de dados, incluindo dados de redes sociais, dados de log de acesso, dados de transações financeiras e outros tipos de dados. Ele pode ser usado para

identificar tendências, padrões e relações entre diferentes tipos de dados, ajudando os usuários a tomar decisões informadas. Além disso, ele pode ser usado para monitoramento de riscos, detecção de fraude e outras finalidades.

4.8.3 *TweetBeaver*

TweetBeaver é uma ferramenta de inteligência artificial (IA) que permite analisar conteúdo de Twitter para produzir provas digitais. Ele usa algoritmos de aprendizado de máquina para analisar tweets, como postagens, comentários, curtidas e outras informações, para identificar tendências, opiniões e sentimentos sobre determinado assunto, rastrear relações entre contas de usuários, identificar conteúdo inapropriado etc. Ele pode ser usado para investigações criminais, monitoramento de atividades maliciosas, análise de risco e outras finalidades.

4.8.4 *MentionMap*

Já o *MentionMap* é outra ferramenta de inteligência artificial (IA) que usa algoritmos de aprendizado de máquina para analisar menções em redes sociais, como postagens, comentários, curtidas e outras informações, para identificar tendências, opiniões, sentimentos e relações entre usuários. Ele pode ser usado para investigações criminais, monitoramento de atividades maliciosas, análise de risco e outras finalidades. Ele permite visualizar as menções em redes sociais em um mapa de relacionamentos, que mostra quem está falando sobre o que e com quem, ajudando a identificar influenciadores e tendências.

4.8.5 *Scraper*

Scraper é uma ferramenta de coleta de dados que permite extrair informações de sites e plataformas na internet, incluindo o Facebook. Ele funciona automatizando a extração de dados de uma página web, coletando dados como postagens, comentários, curtidas, transformando-os em informações, gerando conhecimento.

Para usar um scraper no Facebook, o primeiro passo é acessar o site do Facebook e fazer o login na sua conta. Em seguida, você precisará identificar o conteúdo específico que deseja coletar, como postagens, comentários ou curtidas.

Uma vez que você tenha identificado o conteúdo, você pode configurar o scraper para coletar os dados. Isso pode incluir especificar as palavras-chave para buscar, definir o período para coletar dados e especificar outros parâmetros.

Depois de configurar o scraper, você pode iniciar a coleta de dados. O scraper extrairá as informações do Facebook e as salvará em um arquivo para análise posterior.

5 DA PROTEÇÃO DOS DADOS PESSOAIS E REGULAÇÃO DO USO DE IA

É importante ressaltar que o uso de ferramentas de inteligência artificial para raspagem de dados e outras formas de tratamento de dados precisa observar a legislação, para que não haja violação à proteção de dados pessoais, assegurada pela Constituição Federal, em seu artigo 5º, LXXIX.

A Lei Geral de Proteção de Dados, Lei 13.709/18 estabelece as bases legais para o tratamento de dados pessoais de pessoa física e seus princípios, que devem ser observados sob pena de diversas sanções, dentre as quais multas, publicização da infração, bloqueio e eliminação dos dados pessoais a que se refere a infração na base de dados da empresa, o que afeta em muito a reputação da empresa violadora da lei.

A Lei Geral de Proteção de Dados – LGPD, tem como objetivo *"proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural."*[5] As ferramentas de IA para coleta de dados pessoais do próprio titular é autorizada pelo seu próprio interesse e consentimento, nos termos do art. 7º, I, da LGPD.

Porém, quando o tratamento se referir a dados de terceiros, ele deve estar justificado em uma das bases legais, previstas no mesmo art. 7º a saber: consentimento, obrigação legal, execução de políticas públicas, estudos por órgãos de pesquisa, execução de contrato, defesa em processo judicial, tutela da saúde e proteção à vida, legítimo interesse e proteção ao crédito. Deve, ainda, atender a boa fé e os princípios elencados em seu art. 6º: princípios da finalidade, adequação, necessidade, livre acesso, qualidade, segurança, prevenção, não discriminação, responsabilidade e transparência.

Recentemente, a autoridade de privacidade italiana impôs uma sanção a uma empresa norte-americana alimentada por um sistema de inteligência artificial por violar o Regulamento Geral sobre a Proteção de Dados da União Europeia – GDPR, suspendendo temporariamente seu *chatbot* até a adequação de suas atividades em observância à legislação protetiva dos dados pessoais.

A sanção ocorreu após testes no *chatbot* revelarem riscos concretos para menores de idade e violações do Regulamento 679/2016 (RGPD), incluindo o princípio da transparência. A investigação da autoridade italiana trouxe à tona problemas críticos e riscos para usuários, particularmente menores e indivíduos em situação de vulnerabilidade emocional, decorrentes

[5] Art. 1. Da Lei 13.709/2018.

do *chatbot* com inteligência artificial. Diante das deficiências e questões críticas encontradas, a autoridade italiana de proteção de dados considerou que o tratamento realizado pelo *chatbot* viola os artigos 5, 6, 8, 9 e 25 do RGPD e impôs a restrição provisória do tratamento ao titular, relativamente a todos os utilizadores estabelecidos em território nacional devido à inexistência de qualquer mecanismo de verificação da idade dos usuários. As empresas precisarão continuar buscando um equilíbrio entre verificação de idade e proteção de dados pessoais, considerando as implicações de privacidade da coleta de tais dados. [6]

Existem estudos para regulamentar o uso de inteligência artificial (IA) em vários países e organizações internacionais, incluindo a União Europeia (UE). Em abril de 2021 foi apresentada a Proposta de Regulamento da UE sobre inteligência artificial, que estabelece regras para a utilização da IA em toda a UE, com o objetivo de proteger os direitos fundamentais, garantir a segurança e a privacidade dos cidadãos e fomentar a inovação. A proposta de regulamento proíbe a IA considerada "de alto risco" em algumas áreas, como transporte, saúde e segurança pública, e impõe regras mais rígidas para garantir a transparência e a responsabilidade do uso da IA.

Caso seja aprovado um regulamento geral para uso de IA pela União Europeia, provavelmente ele deverá servir de modelo para o mundo, conforme destacado pelo MIT Techonology Review:

> [...] se a União Europeia conseguir elaborar e adotar um marco regulatório que cumpra satisfatoriamente o objetivo de discernir os principais riscos da aplicação de IA e estabelecer limites e checagens adequados, essa legislação deverá servir como modelo para outros países e regiões. Especialmente nos Estados Unidos, onde se concentram empresas e investimentos de maior porte, presença global, poder e influência na área de IA (como Alphabet/Google e Meta/Facebook), a implementação dessa legislação na Europa deverá fortalecer e ampliar os movimentos já existentes para construção de um modelo regulatório nacional para IA.[7]

Em nível global, a Organização para a Cooperação e Desenvolvimento Econômico (OCDE) publicou em 2019 seus princípios de IA, que incluem a promoção de sistemas justos, transparentes e responsáveis, além de proteger a privacidade e os direitos humanos.[8] Já a UNESCO publicou um documento em 2021, intitulado *"Recomendação sobre a ética da*

[6] GamingTechLaw. Chatbot movido a inteligência artificial banido pela autoridade italiana de privacidade. Publicado em 21/02/23. https://www.gamingtechlaw.com/2023/02/artificial-intelligence-powered-chatbot-italian-privacy-authority/ Acessado em 26/03/2023.

[7] IPEA Centro de Pesquisa em Ciência, Tecnologia e Sociedade**. Lei europeia poderá ser marco global para regulação da inteligência artificial**. Publicado em 02/06/2022. Disponível em: https://www.ipea.gov.br/cts/pt/central-de-conteudo/noticias/noticias/313-lei-europeia-podera-ser-marco-global-para-regulacao-da-inteligencia-artificial Acessado em 29/03/23.

[8] Migalhas. Inteligência Artificial, princípios e recomendações da OCDE. Publicado em 09/07/22. Disponível em: Inteligência Artificial, princípios e recomendações da OCDE. Acessado em 29/03/23.

inteligência artificial", que orienta os países e organizações a desenvolverem políticas e práticas de IA éticas e responsáveis, promovendo a transparência, a diversidade e a proteção dos direitos humanos.[9]

No Brasil, ainda não existe uma legislação específica para regulamentar o uso da IA, mas existem algumas iniciativas em andamento, como o projeto de lei 21/2020, que visa criar regras para o uso da IA em diferentes setores, estabelecendo fundamentos, princípios e diretrizes para o desenvolvimento e a aplicação da IA no Brasil, conforme destaque da Agência da Câmara de notícias:

> Entre outros pontos, a proposta estabelece que o uso da IA terá como fundamento o respeito aos direitos humanos e aos valores democráticos, a igualdade, a não discriminação, a pluralidade, a livre iniciativa e a privacidade de dados.[10]

6 CONCLUSÃO

A inteligência artificial para extração de dados digitais para produção de provas em processos judiciais pode ser primordial para garantia de direitos fundamentais como o direito à honra, a não discriminação, à saúde e à liberdade, dentre outros.

A IA pode ser usada para identificar autores de postagens em mídias sociais, analisar metadados de fotos digitais, pesquisar patrimônio de devedores em processos de execução e analisar mensagens de texto em investigações criminais.

As vantagens da IA incluem a capacidade de processar grandes quantidades de dados rapidamente e identificar padrões e tendências. No entanto, a volatilidade dos dados digitais exige que os profissionais do direito tenham conhecimento ou auxílio de profissionais da área de computação forense e ferramentas de IA para coletar e armazenar dados de forma válida.

É importante notar que, como com qualquer tecnologia, é necessário considerar questões éticas e legais ao usar ferramentas de inteligência artificial e ciência de dados no direito, para que não haja violação ao direito da privacidade. Isso inclui garantir que as ferramentas sejam usadas de forma ética e legal, bem como garantir que as decisões baseadas em dados sejam justas e imparciais. Além disso, é importante considerar que essas tecnologias não devem

[9] UNESCO. Ética da Inteligência Artificial (IA) no Brasil. Atualizado em 15/03/23. Disponível em: https://www.unesco.org/pt/fieldoffice/brasilia/expertise/artificial-intelligence-brazil#:~:text=Em%20novembro%20de%202021%2C%20a,este%20tema%20na%20sociedade%20brasileira. Acessado em 29/03/23.

[10] Agência Câmara de Notícias. Projeto cria marco legal para uso de inteligência artificial no Brasil. Publicado em 04/03/2020. Disponível em: https://www.camara.leg.br/noticias/641927-projeto-cria-marco-legal-para-uso-de-inteligencia-artificial-no-brasil/ Acesso em: 27/03/23.

substituir completamente a sabedoria humana, mas sim ser usadas como uma ferramenta a mais para auxiliar e melhorar o processo de tomada de decisão.

O uso de ferramentas de inteligência artificial (IA) na coleta de dados digitais para comprovação de fatos em processos judiciais oferece diversas vantagens, incluindo eficiência, precisão, identificação de informações relevantes, automação de tarefas repetitivas, armazenamento seguro e fundamentação para tomada de decisões:

Eficiência: A IA é capaz de processar grandes volumes de dados digitais rapidamente, tornando o processo de coleta de informações muito mais eficiente do que seria se realizado manualmente.

Precisão: A IA pode ser programada para encontrar padrões e tendências específicas nos dados que podem não ser facilmente detectáveis por humanos, aumentando a precisão e a confiabilidade da prova.

Identificação de informações relevantes: A IA pode ser usada para identificar informações relevantes em grandes conjuntos de dados, como mensagens de texto ou posts em redes sociais, ajudando os advogados a encontrar evidências que possam ser usadas em processos judiciais.

Automação de tarefas repetitivas: As ferramentas de IA podem automatizar tarefas repetitivas, como a pesquisa de jurisprudência, liberando os advogados, juízes, procuradores, servidores e outros profissionais da área jurídica para trabalhos mais complexos e de maior valor agregado.

Armazenamento seguro: A IA pode ser usada para armazenar dados digitais coletados de forma segura, garantindo que eles possam ser usados como prova em processos judiciais sem risco de perda ou corrupção, garantindo a integridade da prova.

Tomada de decisão mais informada: A IA pode ser usada para analisar os dados coletados e ajudar os profissionais do direito a tomar decisões fundamentadas com base em insights e padrões encontrados nos dados.

Conforme exposto, o uso das ferramentas de IA, de forma ética e legal, oferece vantagens significativas na coleta e preservação de dados digitais para comprovação de fatos em processos judiciais, aumentando a confiabilidade da prova e a segurança jurídica, contribuindo com a proteção dos direitos fundamentais, em observação ao Objetivo de

Desenvolvimento Sustentável nº 16 da Organização das Nações Unidas, de promover a paz e a justiça.

REFERÊNCIAS

Agência Câmara de Notícias. **Projeto cria marco legal para uso de inteligência artificial no Brasil.** Publicado em 04/03/2020. Disponível em: https://www.camara.leg.br/noticias/641927-projeto-cria-marco-legal-para-uso-de-inteligencia-artificial-no-brasil/ Acesso em: 27/03/23.

Artificial Intelligence and Blockchain in Digital Forensics (River Publishers Series in Digital Security and Forensics) (p. 246). River Publishers. Edição do Kindle. Tradução livre: "The technical process of identifying, preserving, collecting, and presenting digital evidence so that it becomes applicable to the case of law is known as digital forensics. In reality, any information stored or retrieved from digital technology might be considered a piece of electronic evidence that can be analysed throughout a digital forensics' inquiry."

ANTUNES, Mário. RODRIGUES, Baltazar. **Introdução a Cibersegurança. A Internet, Os Aspectos Legais e a Análise Digital Forense.** Editora FCA. Lisboa, 2018, p. 30.

AWS Academy **Machine Learning Foundations.** Module 02 Student Guide.

Analisador de cabeçalho (header) de e-mail - kingHost. Disponível em: https://King.host/wiki/ferramenta-header/ Acesso em: 26/03/23.

BAYER, Judy e TAILLARD, Marie. Story-driven data analysis. From Data to action. **A Harvard Business Review Insight Center Report.**

BELLUZZO, Regina Celia Baptista, VALENTE, Vânia Cristina Pires Nogueira. **A Competência em informação, as competências digitais e o protagonismo dos agentes sociais e mediadores na sociedade contemporânea.** Competencias em información y transformación digital de la sociedade. VALERO, Pablo Parra, CUEVAS-CERVERÓ Aurora, SIMEÃO Elmira, RUIZ, Maria Jesús Colmenero (Coordinadores). Universidad Complutense de Madrid Facultad de Ciencias de la Documentación, Departamento de Biblioteconomía y Documentación, 2021. Disponível em: https://eprints.ucm.es/id/eprint/71169/ Acesso em: 27/03/2023.

BRASIL. **Constituição Da República Federativa Do Brasil De 1988**. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm Acesso em: 05 fevereiro 2023.

BRASIL. **Lei 13.709, de 14 de Agosto de 2018 (Lei Geral de Proteção de Dados Pessoais)** Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm Acesso em: 29/03/23.

CASELLI, Guilherme. **Manual de Investigação Digital.** São Paulo: Editora JusPodivm, 2022, p. 41.

Digital 2023 July Global Statshot Report (Janeiro 2023). Publicado em 20/02/23. Disponível em: https://www.slideshare.net/DataReportal/digital-2023-global-overview-report-summary-version-january-2023-v02 Acesso em: 27/03/23.

DAVENPORT, Tom. What to ask your "numbers people". From Data to action. **A Harvard Business Review Insight Center Report, p. 2.**

ESCOVEDO, Tatiana; KOSHIYAMA, Adriano. **Introdução a Data Science.** Casa do Código. Edição do Kindle.

IPEA Centro de Pesquisa em Ciência, Tecnologia e Sociedade**. Lei europeia poderá ser marco global para regulação da inteligência artificial**. Publicado em 02/06/2022. Disponível em: https://www.ipea.gov.br/cts/pt/central-de-conteudo/noticias/noticias/313-lei-europeia-podera-ser-marco-global-para-regulacao-da-inteligencia-artificial Acessado em 29/03/23.

MARTINS, João Marques. Revista da Faculdade de Direito da Universidade de Lisboa. Número Temático: Tecnologia e Direito, 2022. **Inteligência Artificial e Direito: Uma Brevíssima Introdução**. Disponível em: https://www.fd.ulisboa.pt/wp-content/uploads/2022/12/Mariana-Pinto-Ramos.pdf Acesso em: 27/03/23.

Migalhas. **Inteligência Artificial, princípios e recomendações da OCDE.** Publicado em 09/07/22. Disponível em: Inteligência Artificial, princípios e recomendações da OCDE. Acessado em 29/03/23.

MENON, Sunand. Stop assuming your data will bring you riches. From data to action. **A harvard business review insight center report.**

MORAIS, Izabelly Soares de... [et al.]. **Introdução a Big Data e Internet das Coisas (IoT)** [recurso eletrônico]; [revisão técnica:]. – Porto Alegre: SAGAH, 2018

NUNES, Marcelo Guedes. **Jurimetria: como a estatística pode reinventar o direito**. São Paulo: Thomson Reuters Brasil, 2020. Edição do Kindle.

O'CONNELL, Andrew e FRICK, Walter. But what does it mean? Welcome to "from data to action". From Data to action. **A Harvard Business Review Insight Center Report.**

OETTINGER, William. **Aprenda Computação Forense**. São Paulo: Novatec Editora Ltda., 2021. Edição do Kindle.

ONLINE COMPUTER LIBRARY CENTER (2002). **Preservation metadata and the OAIS information model: a metadata framework to support the preservation of digital objects** (PDF). Ohio, USA: OCLC. 51 páginas. Acesso em: 28 de novembro de 2019, apud Tribunal Regional do Trabalho da 7ª Região. MSCiv 0080297-42.2021.5.07.0000. Desembargadora Maria Roseli Mendes Alencar. 10/06/2021. Disponível em: file:///C:/Users/Global/Downloads/Documento_5b9bc6e.pdf. Acesso em: 27/03/23.

PINHEIRO, Patricia Peck. **Direito Digital**. São Paulo: Saraiva Educação, 2021.

RAJASEKAR, Vani, SATHYA, K, VELLIANGIRI, S e KARTHIKEYAN, P. **Blockchain-based Identity Managemente Systems in Digital Forensics**. Artificial Intelligence and Blockchain in Digital Forensics (River Publishers Series in Digital Security and Forensics). River Publishers, 2023. Edição do Kindle.

RUSSELL, Stuart J.; Norvig, Peter. **Inteligência Artificial - Uma Abordagem Moderna**. Rio de Janeiro: GEN Grupo Editorial Nacional S.A. Publicado pelo selo LTC Livros Técnicos e Científicos Ltda., 2022. Edição do Kindle.

SAXENA, Ishi, USHA, G, VINOTH, N.A.S., VEENA, S. e NANCY, Maria. **The Future of Artificial Intelligence in Digital Forensics: A Revolutionary Approach**. Artificial Intelligence and Blockchain in Digital Forensics (River Publishers Series in Digital Security and Forensics) River Publishers, 2023. Edição do Kindle.

SHARMA, Hemant. **O que é ciência de dados? Um guia para iniciantes em Ciência de Dados.** Acesso em: https://www.edureka.co/blog/what-is-data-science/

STATON, J.M. (2012). **Introduction to Data Science**, Third Edition. iTunes Open Source eBook. Available: https://itunes.apple.com/us/book/introduction-to-data-science/id529088127?mt=11

THAMAY, Rennan e TAMER, Mauricio. **Provas no Direito Digital: conceito da prova digital, procedimentos e provas digitais em espécie**. São Paulo: Thomson Reuters Brasil, 2020.

THE GUARDIAN. **NSA FILES: DECODED. What revelations mean for you**. 2013. https://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded#section/2, p. 2, Acesso em: 26/03/2023

The OCLC/RLG Working Group on Preservation Metadata. **Preservation Metadata and the OAIS Information Model. A Metadata Framework to Support the Preservation of Digital Objects.** Junho/2002. Disponível em: https://www.oclc.org/content/dam/research/activities/pmwg/pm_framework.pdf Acesso em: 27/03/23.

UNESCO. **Ética da Inteligência Artificial (IA) no Brasil.** Atualizdo em 15/03/23. Disponível em: https://www.unesco.org/pt/fieldoffice/brasilia/expertise/artificial-intelligence-brazil#:~:text=Em%20novembro%20de%202021%2C%20a,este%20tema%20na%20sociedade%20brasileira. Acessado em 29/03/23.

UNIÃO EUROPEIA. **Proposta de Regulamento (EU) 2021/0106 do Parlamento Europeu e do Conselho. (Regulamento Inteligência Artificial).** Publicado em 21/04/2021. Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52021PC0206&from=EN. Acesso em: 29/03/23.

UNIÃO EUROPEIA. **Regulamento (EU) 2016/679 do Parlamento Europeu e do Conselho. (Regulamento Geral sobre a Proteção de Dados)**. Publicado em 27/04/2016. Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32016R0679 Acesso em: 29/03/23.