

INTELIGÊNCIA ARTIFICIAL NO PODER JUDICIÁRIO: LIÇÕES DO PROJETO VICTOR

ARTIFICIAL INTELLIGENCE IN JUDICIARY POWER: LESSONS FROM VICTOR PROJECT

Tainá Aguiar Junquillo 1
Mamede Said Maia Filho 2

Resumo: O Projeto Victor, executado pelo Poder Judiciário brasileiro, apresenta-se como o primeiro a aplicar Inteligência Artificial em Cortes Constitucionais no mundo. Este artigo tem por objetivo o estudo de caso do Projeto, visando apresentar os resultados obtidos e expor a metodologia de desenvolvimento utilizada. A intenção é, a partir dos dados da pesquisa, possibilitar a realização de diagnósticos futuros e contribuir para a transparência e o controle democrático – fundamentais no delineamento das políticas institucionais – de iniciativas como o Projeto Victor, que envolvem o uso de Inteligência Artificial pelo Poder Público, no que vem sendo chamado de Artificial Intelligence for Social Good (AI4SG).

Palavras-chave: Inteligência Artificial. Projeto Victor. Aprendizagem de Máquina.

Abstract: Victor Project is presented as the first Artificial Intelligence project applied to Constitutional Courts in the world, conducted by the Brazilian Judiciary Power. This article aims to introduce a case study of this project, presenting the achieved results and showcasing the development methodology used on the research. The intention is, through the data, to allow future analysis by researchers and to contribute to the transparency and democratic control of the Victor Project, which are essential in establishing institutional policies regarding the use of Artificial Intelligence by Public Authorities, in what has been referred as Artificial Intelligence for Social Good (AI4SG).

Keywords: Artificial Intelligence. Project Victor. Machine Learning.

Doutoranda em Direito, Estado e Constituição pela Universidade de Brasília (UnB); Mestre em Direito Processual pela Universidade Federal do Espírito Santo (UFES).
Lattes: <http://lattes.cnpq.br/5848504606151120>.
ORCID: <https://orcid.org/0000-0001-8638-0670>.
E-mail: taina.aguiarj@gmail.com

Doutor em Direito, Estado e Constituição pela Universidade de Brasília (UnB).
Lattes: <http://lattes.cnpq.br/6611037107398589>.
ORCID: <https://orcid.org/0000-0002-9191-6430>.
E-mail: mamedesaid@gmail.com

Introdução

O avanço da Tecnologia da Informação tem afetado e transformado as mais diversas áreas do conhecimento humano e da sociedade, impactando, significativamente, a entrega de dados e informações a um amplo leque de indivíduos e organizações. De fato, as “possibilidades cada vez maiores de receber, armazenar e gerenciar imensas quantidades de dados de diversas fontes [...] levaram a tendências recentes comumente chamadas de *big data* e *data mining*.” (HULDTGREN, 2015, p. 762).

Nesse sentido, Floridi (2014) denomina esse momento da evolução de revolução informacional, diante do qual se conclui que existem outros seres informacionais, além do ser humano, os quais o autor chama de “Inforgs”. O contexto da internet das coisas (IoT) viabiliza a produção em grande escala de dados, que são coletados e armazenados nos dispositivos cotidianos utilizados pelos cidadãos e que servem para os mais variados usos (MAGRANI, 2018).

Uma das tecnologias que mais evoluiu na Quarta Revolução Industrial (ou Indústria 4.0) é a denominada Inteligência Artificial (IA) (SCHWAB, 2019). As previsões de que as mudanças promovidas pela IA atingirão sobremaneira o setor de serviços, o mercado e trabalho e até o modo como vivemos têm, cada vez mais, se tornado realidade (SCHWAB, 2019).

O aprendizado de máquina (AM) tem como premissa a ideia de que um computador pode aprender uma tarefa determinada por meio de um conjunto de dados, que servem como exemplos para seu treinamento (LOURIDAS, 2016). No desenvolvimento de projetos técnico-científicos que pretendem viabilizar a aplicação de aprendizado de máquina há, portanto, uma condição *sine qua non*: a necessidade de utilização de um grande volume de dados (PASQUALE, 2015). Em projetos desenvolvidos na área do Direito, essa ocorrência não é diferente.

O elevado número de processos que tramitam no Poder Judiciário brasileiro, atualmente, não comporta mais um tratamento feito pela programação convencional. O adequado armazenamento e tratamento dos documentos, hoje, são gerados por múltiplas fontes e realizados por técnicas de *machine learning*. Esta evolução evidencia a eficiência do uso de algoritmos estatísticos, a partir de uma robusta e numerosa base de dados (GIUFFRIDA, 2018).

É nesse contexto de incremento da tecnologia na coleta e de processamento e disseminação da informação no Poder Judiciário que surgiu o Projeto Victor, fruto da parceria firmada pelo Supremo Tribunal Federal (STF) com a Universidade de Brasília (UnB). O Projeto é uma ferramenta de utilização de IA capaz de identificar, nos processos que chegam ao STF para realização do controle de constitucionalidade difuso, a presença do requisito da repercussão geral (RG) exigido pelo artigo 102, § 3º, da Constituição Federal, e pelo artigo 1.035 do Código de Processo Civil.

Este artigo pretende apresentar e detalhar como se deu o processo de desenvolvimento do Projeto Victor, contextualizando suas fases. Também busca explicitar quais os principais temas que chegaram ao STF nos dois anos anteriores à sua execução (entre 01/04/2017 e 30/04/2019), analisando a metodologia de extração de dados para os temas e o que esta pode trazer como contribuição para o Direito e para a tecnologia.

Realizou-se, portanto, um estudo de caso, metodologia investigativa que permite compreender e aprofundar os elementos e aspectos que permearam o desenvolvimento deste Projeto (LAVILLE; DIONNE, 1999). Assim, visou-se identificar as etapas, as dificuldades e os dados obtidos, em atenção à transparência que deve pautar o uso de Inteligência Artificial, principalmente pelo Poder Público.

A problemática deste artigo envolve um projeto que se insere no contexto da chamada AI4SG (*Artificial Intelligence for Social Good*), ou seja, Inteligência Artificial para o Bem Social. Este conceito sinaliza a tendência mundial de aplicação de soluções tecnológicas em IA para resolução de problemas sociais (FLORIDI *et al.*, 2020). Como observa Russell (2019, p. 11), “as máquinas são benéficas na medida em que se espera que suas ações atinjam nossos objetivos”; no caso em análise, essa premissa se estende, desde que a tecnologia esteja a serviço da sociedade e dos impulsos em prol de seu dinamismo.

Quando os usos tecnológicos se dão por iniciativa do Poder Público, recomendam-se estudos como este, que exponham, de forma transparente, os dados e resultados alcançados. Afinal, para que possamos, enquanto sociedade, confiar no produto da IA, é necessá-

rio que os modelos de aprendizado de máquina sejam transparentes e interpretáveis (GHANI, 2016). Tanto os modelos extraídos a partir de conjuntos de dados como os algoritmos de aprendizado de máquina devem ser suficientemente compreensíveis. Tal compreensão irá garantir que os processos associados às tomadas de decisão automatizadas sejam confiáveis (*trustworthy AI*) e que os sistemas de IA sejam centrados no ser humano (*human-centric AI*), beneficiando o crescimento inclusivo e o bem-estar social.

A resposta à problematização é, portanto, a exposição e a análise do modelo utilizado em uma relevante e inovadora experiência, a qual pode contribuir para o debate, ensino e desenvolvimento de novas ferramentas desse tipo nas tecnologias de computação cognitiva. Além disso, almeja-se expor e difundir dados capazes de formar importante diagnóstico dos temas de repercussão geral atualmente analisados pela Corte Suprema brasileira.

Por que o Brasil foi pioneiro em aplicar um projeto de IA a uma Corte Constitucional?

A Corte Constitucional brasileira é altamente sobrecarregada. Mesmo com o Tribunal registrando, em 2020, uma queda, na comparação com 2019, do número de processos recebidos, a quantidade naquele ano foi de 75.137 processos (SUPREMO TRIBUNAL FEDERAL, 2021, p. 26). Apenas entre recursos extraordinários (RE), recursos extraordinários com agravo (ARE) e agravos de instrumento (AI), foram recebidos 49.783 recursos. O ano de 2020 foi marcado, ainda, pelo aumento expressivo do número de temas de repercussão geral com mérito julgado, com positivos impactos em todo o sistema jurídico nacional. A ampliação de julgamentos de processos com repercussão geral reconhecida em 2020 altera a preocupante relação entre os temas aguardando julgamento e a capacidade de julgamento até então presente no Supremo Tribunal Federal (SUPREMO TRIBUNAL FEDERAL, 2021, p. 50), para o que, certamente, contribuiu a implantação do Projeto Vítor.

Seja qual for o ângulo de análise, o número de processos em tramitação no STF é infinitamente superior, em termos comparativos, aos de cortes constitucionais como a dos Estados Unidos e a da Alemanha. Esses países estabelecem um filtro de relevância para que apenas os processos mais significativos tenham seu mérito apreciado, de modo compatível com a capacidade dos tribunais para julgá-los. Nos Estados Unidos, cuja Suprema Corte recebe cerca de 7 mil casos por ano, não mais que 100 deles são decididos pelo plenário (MEADOR; MITCHELL, 2009, p. 25). Na Alemanha, o número também é similar, e a maior parte é considerada sem significação constitucional (FEDERAL CONSTITUTIONAL COURT, 2016).

Como dito, o aprendizado de máquina depende de grande massa de dados, uma vez que a máquina aprende com informações do passado (DOMINGOS, 2012). Esse expressivo aumento de dados produzidos e armazenados diariamente no Tribunal tornou, portanto, o uso de *machine learning* uma alternativa adequada para o aperfeiçoamento do sistema de repercussão geral e da própria jurisdição constitucional. Isso é, o fato de o sistema judiciário brasileiro, em especial do STF, ser congestionado, implica no surgimento diário de uma grande massa de dados, o que possibilitou que o Brasil fosse pioneiro na aplicação de IA a uma Corte Constitucional.

Com o escopo de tornar a jurisdição constitucional mais célere e eficiente, a gestão da Presidente Ministra Carmen Lúcia decidiu, então, iniciar o projeto de IA pioneiro no Judiciário brasileiro. Cabe ressaltar que iniciativas como essa podem ser classificadas como AI4SG, isso é, como IA que reflete importantes benefícios para a sociedade, ao cumprirem seu objetivo “na medida em que ajudam a reduzir, mitigar ou erradicar um determinado problema de significado moral” (HAGER *et al.*, 2017, p. 24). Assim, dizem respeito à

[...] implantação de sistemas de IA de maneira que (i) previnam, mitiguem ou resolvam problemas que afetam adversamente a vida humana e/ou o bem-estar do mundo natural, e/ou (ii) possibilitem desenvolvimentos socialmente preferíveis e/ou

ambientalmente sustentáveis (FLORIDI *et al.*, 2020, p. 1773).

Nesse sentido, o Projeto Victor pretendeu aplicar IA ao STF para melhorar a prestação jurisdicional, conferindo eficiência ao controle de constitucionalidade. Tal iniciativa não pode, entretanto, perder de vista os princípios da transparência e a contextualização de seu desenvolvimento, o qual será apresentado a seguir.

Fases de desenvolvimento do Projeto Victor: o que cada uma ensinou

As tecnologias de aprendizado de máquina podem ser aplicadas em várias áreas da vida humana, sendo particularmente benéficas quando contribuem para a promoção do bem-estar social. Para que a IA se concretize como AI4SG, Floridi *et al.* (2020) sugerem alguns fatores a serem cumpridos, tais como a criação de um produto de interação amigável com os possíveis usuários e transparência de propósitos e resultados dos projetos. Ora, é fundamental que toda proposta de criação de IA seja humanocêntrica, ou seja, criada para atender aos interesses da coletividade humana. A partir dessa compreensão, este artigo narra como foi o andamento do Projeto de Pesquisa & Desenvolvimento de aprendizado de máquina (*machine learning*) sobre dados judiciais das repercussões gerais do Supremo Tribunal Federal – (STF), conhecido com Projeto Victor. Além disso, aponta conclusões relevantes, com o intuito de promover sua transparência.

É importante destacar, inicialmente, o caráter multidisciplinar no desenvolvimento do Projeto Victor, que envolveu pesquisadores da Faculdade de Direito e dos cursos de Engenharia de *Software* e Ciência da Computação da UnB, trabalhando em conjunto na análise, tratamento e utilização de dados. Tal fato em si representa uma iniciativa inovadora, por abranger alunos e professores de áreas tão diversas do conhecimento, o que propiciou um debate interdisciplinar e, portanto, mais profundo, capitaneado pela Universidade, em parceria com o órgão de cúpula do Poder Judiciário. O Projeto se desenvolveu em três fases gerais, com cada uma apresentando desafios, dados e perspectivas interessantes para o aprendizado, a interpretação e a evolução da utilização de Inteligência Artificial, o que certamente serve de lição para futuras aplicações de IA ao âmbito jurídico.

Na fase inicial do Projeto Victor, o STF apresentou aos pesquisadores a necessidade de criação de uma ferramenta que realizasse a separação das peças jurídicas que compõem os processos que chegam ao Tribunal. Diante dessa demanda inaugural, as equipes se concentraram no desenvolvimento de um separador de peças, que facilitasse o trabalho cotidiano dos mais diversos segmentos da Corte que atuam com a questão.

Nesse estágio, a massa de dados apresentada pelo STF continha processos de diversas origens e tribunais, o que gerou alguns contratempos na identificação de seus enunciados. A maior dificuldade encontrada teve relação com os processos advindos de plataformas eletrônicas diversas, ou que haviam sido obtidos por meio de digitalização de pleitos originariamente “físicos”. Tais processos apresentavam caracteres de difícil compreensão e, por isso, tiveram que passar pelo reconhecimento óptico de caracteres (OCR – *Optical Character Recognition*) para que houvesse a transformação das imagens em textos (PEREZ *et al.*, 2017). A adoção do método foi possível porque os textos que passam pelo procedimento de OCR, mesmo que não fiquem perfeitos, trazem um resultado bastante preciso e confiável depois da estruturação e limpeza dos dados.

O primeiro desafio do Projeto surgiu, então, em virtude da deficiência na uniformização do processo eletrônico no país, que é um problema para a maioria dos projetos de desenvolvimento que pretendem aplicar IA ao Judiciário. Providência dessa natureza (OCR) torna-se, em geral, necessária, em virtude do grande número de plataformas de processos eletrônicos ora em operação no país. Este é um entrave a ser vencido para que se alcance um adequado padrão de comunicação entre os sistemas processuais informatizados (CUEVAS, 2021).

Existem vários sistemas em uso nos tribunais hoje (PJe, Projudi, e-SAJ, e-Proc, Tucujuris e Apolo), com algumas Cortes tendo, inclusive, mais de um sistema funcionando simultaneamente. Essas plataformas variadas, por não guardarem uniformidade entre elas, prejudicam o

fluxo de peticionamento e o acompanhamento processual. Por isso, uma percepção empírica inaugural e importante do Projeto Victor relaciona-se à forma discrepante como, atualmente, tramitam os processos em cada Tribunal. Este é um fator que, além de limitar a eficácia da prestação jurisdicional, também representa um entrave à aplicação de projetos tecnológicos ao Judiciário.

Essa primeira conclusão do Projeto leva à compreensão de que, se o Brasil quer despontar como produtor de ambientes de inovação, em especial na área jurídica, deve ter como passo inicial a realização de investimentos para facilitar a informatização e estruturação dos dados processuais. O desenvolvimento desta pesquisa reafirmou na prática, portanto, a compreensão teórica de que o incremento de utilização das novas tecnologias está intimamente ligado ao avanço e realização de políticas públicas de melhoria na qualidade do acesso à informação. Afinal, acesso à informação é também uma forma de concretização do acesso à justiça pelos cidadãos e de promoção da transparência (BECKER, 2019).

Cabe destacar que o CNJ tem adotado importantes iniciativas para a digitalização da justiça e melhoria no sistema informacional do Poder Judiciário. Merece menção, em especial: a) a Resolução nº 335, que criou a Plataforma Digital do Poder Judiciário Brasileiro – (PDPIJ-Br); b) a Resolução nº 345, que criou o chamado “Juízo 100% Digital”; e c) a Resolução nº 331, que estabeleceu a Base Nacional de Dados do Poder Judiciário – (DataJud) – como fonte primária de dados do Sistema de Estatística do Poder Judiciário (SIESPJ) para os tribunais indicados nos incisos II a VII do artigo 92 da Constituição Federal.

Procedimentos e iniciativas como essas são relevantes para que se estabeleça a integração entre os sistemas existentes, de modo a padronizar a linguagem e o trânsito que envolve as diversas fases e matérias processuais. Cumprem, ainda, o importante papel de fomentar e facilitar a criação de projetos de desenvolvimento e a aplicação de novas tecnologias de automação, tornando o sistema judicial também mais célere e eficaz.

A segunda fase do Projeto Victor nasceu da observação das atividades desenvolvidas pelo STF no julgamento dos recursos extraordinários. Os pesquisadores da área jurídica, por meio da observação do trabalho dos servidores do STF que atuavam no recebimento dos Recursos Extraordinários, concluíram que eram priorizadas cinco peças na análise dos processos referentes a temas cuja repercussão geral tenha sido reconhecida: o acórdão, o recurso extraordinário, o agravo de recurso extraordinário, o despacho e a sentença.

Foram, então, adotadas três etapas para a solução do problema de separação de peças: 1) criação de um conjunto verdade para treinamento de modelos; 2) análise exploratória sobre o conjunto de dados; 3) “arquitetagem” de modelos de AM (aprendizado de máquina) para classificação de peças (SILVA, 2018, p. 90).

O desenvolvimento do separador de peças permitiu, dessa forma, um refinamento dos dados coletados, de modo a trabalhar um espaço amostral no qual se verificava aqueles que eram similares e dos quais se obtinha uma classificação mais qualitativa. Cabe ressaltar que essa etapa, embora não estivesse prevista no escopo inicial do Projeto, apresentou-se como uma demanda urgente do Tribunal. Sua concretização e disponibilização de resultados cumpriu um importante papel na otimização do tempo de trabalho, para identificação dos processos e sua respectiva análise.

Após a entrega do separador de peças, feitas a limpeza e a organização dos dados, partiu-se para a terceira fase, foco de desenvolvimento do escopo principal do Projeto: a criação de tecnologia capaz de reconhecer a qual tema de repercussão geral o processo que chega ao STF está relacionado. Para tanto, foram escolhidos os 27 temas de RG mais frequentes na Corte, os quais apresentavam, por essa razão, maior base de dados a serem trabalhados. Ou seja, no estágio inicial, foram selecionados aqueles temas que, durante o prazo de dois anos (2017-2019), apareceram com maior repetição no STF. Esses temas tinham, portanto, um número mais significativo de processos relacionados, pré-selecionados e já classificados pela equipe do Tribunal.

A equipe do Direito realizou a divisão dos 27 temas em seis classes, por assuntos correlatos: a) classe de temas mais frequentes em Processo Civil; b) classe de temas mais frequentes de Juizados Especiais; c) classe de temas mais frequentes afetos à área de Fazenda Pública; d)

classe de temas mais frequentes de direito previdenciário; e) classe de temas mais frequentes relacionados a instituições bancárias; e f) classe de temas mais frequentes em danos morais. A partir dos dados (processos enviados pelo STF e já classificados em algum dos temas de RG), os especialistas em Direito da equipe iniciaram um trabalho chamado de *tagueamento* (cadastro e rastreamento dos processos analisados) das peças, para realização do aprendizado supervisionado de máquina.

Dentro do Projeto Victor, o *tagueamento* consistia no trabalho da equipe, que avaliou quais eram as palavras-chave fundamentais presentes nas peças processuais que, em sua visão, serviam para identificar e classificar o processo em determinado tema. Assim, a partir da leitura e releitura dos processos que já haviam sido anteriormente classificados pela equipe do STF em um tema, tentou-se encontrar quais termos presentes nas peças serviam para identificar e caracterizar cada processo dentro de um assunto da repercussão geral.

Essa metodologia para extração de dados em pesquisas de desenvolvimento de IA consiste na realização de perguntas formais, tal como “o que se pode excluir e/ou incluir, pensando na relevância do conteúdo daquele documento, para identificação dele como relativo a um determinado tema e para consecução do objetivo?” Esse trabalho de análise de metadados levou a um agrupamento, ou seja, formou-se um conjunto de palavras-chave escolhidas pelos especialistas e consideradas como representativas daquele tema de RG, constituindo uma base de dados bastante consistente.

Em seguida, esses dados foram apresentados à equipe de tecnologia e comparados aos dados obtidos pela programação. Nesse ponto, a equipe de tecnologia identificou que a definição de RG nos processos jurídicos apresentados tratava de questões que estavam associadas a múltiplos rótulos, requerendo, assim, métodos de aprendizado *multi-label* ou multirrótulo. Esse desafio demandou da equipe de tecnologia algumas considerações e aplicação de método.

O Direito não é uma ciência exata, é argumentação e construção linguística de *topoi*, isso é, dá-se pela construção linguística de lugares comuns (VIEHWEG, 1979). Analisando-se por esse prisma, percebeu-se que alguns processos que chegam ao STF não são de fácil identificação, podendo se enquadrar, inclusive, em mais de um tema. Coube à equipe de tecnologia, então, escolher métodos que identificassem probabilidades de o processo pertencer a cada um dos 27 temas selecionados.

Porém, os processos que tratam de assuntos específicos, como, por exemplo, os relativos aos direitos dos servidores federais às diferenças em seus vencimentos, após a mudança para o regime estatutário (tema 951), dificilmente são classificados em conjunto com outros temas. Ou seja, são mais fáceis de serem identificados por dizerem respeito a áreas especiais e/ou por tratarem de situações reguladas em leis que restringem sua aplicação a contextos fáticos e a circunstâncias bastante singulares.

No que tange aos recursos extraordinários que abordam teses genéricas, como, por exemplo, o tema que envolve a fundamentação das decisões judiciais (tema 339) e o que compreende a não observância do contraditório e da ampla defesa (tema 660), em geral, é difícil que o STF os classifique em apenas um enunciado. Isso equivale a dizer que, usualmente, nesses casos, vários temas de RG podem se encaixar em um mesmo RE. Nesse sentido, apresenta-se a questão dos processos que tangenciam temas associados ao processo de rotulação “multirrótulo” (que comportam, portanto, várias classificações), os quais exigiram da equipe de tecnologia a adoção de metodologias especiais que geraram, ao final, resultados bastante precisos.

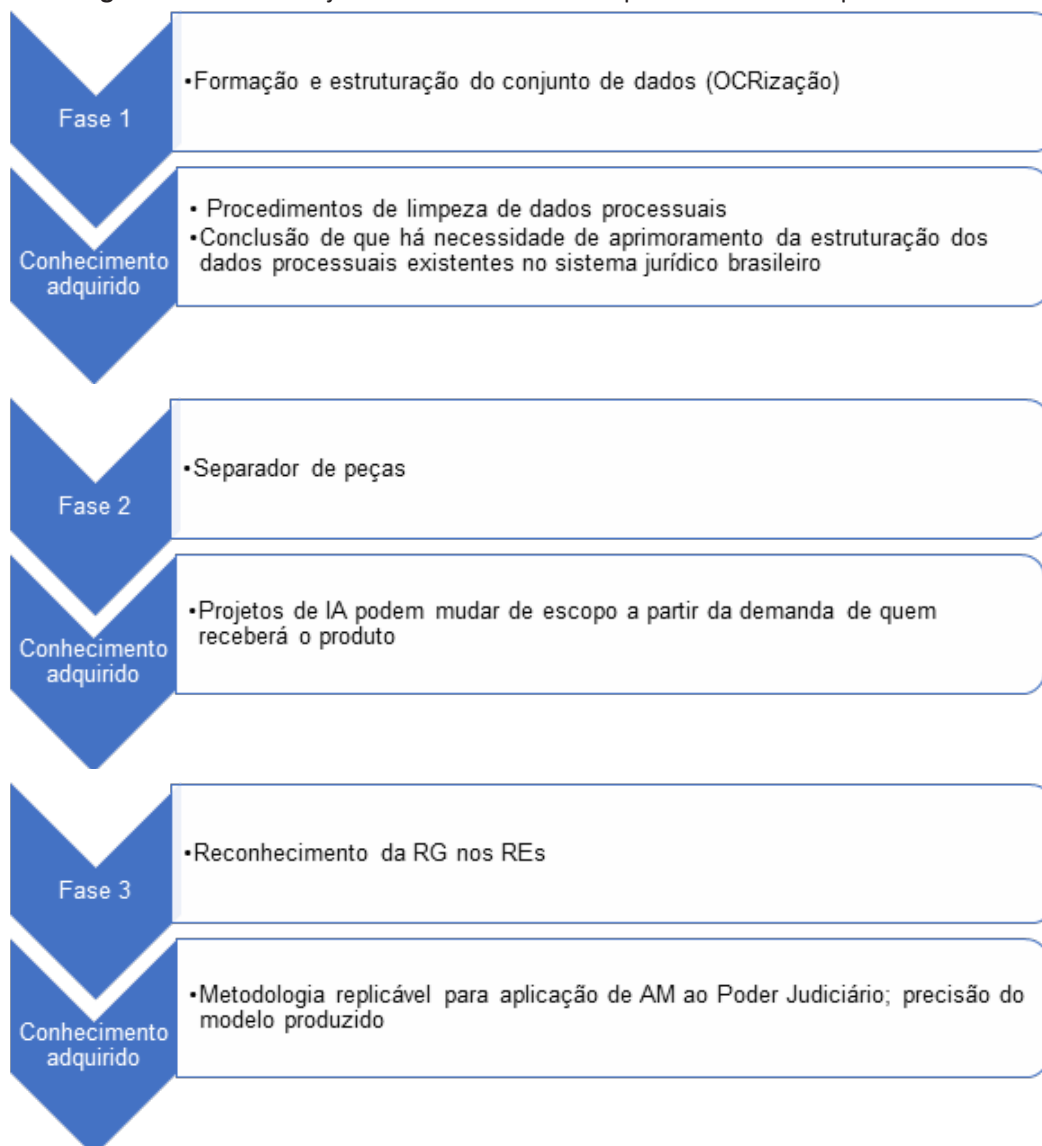
Em relação a esses resultados, quando se pergunta pela precisão, um termo muito utilizado é “acurácia”, que significa saber se o “resultado da medição é preciso, ou seja, em que medida corresponde ao estado de coisas ‘real’ no mundo (ao valor ‘real’ no caso de medições quantitativas)” (KROES; POE, 2015, p. 166). Trata-se de um conceito que remete à exatidão (precisão geral), indicando o grau de concordância que há entre o resultado da medição e o dito valor verdadeiro da grandeza. É entendido como verdadeiro aquele valor que é aceito porque foi estabelecido por uma definição ou consenso.

A medição, nesta pesquisa, foi obtida utilizando-se a métrica *F1 Score* – valor entre 0

(0%) e 1 (100%), tendo sido alcançado, para os temas selecionados, o valor de 0,927, o que revela que a precisão do resultado chegou muito próximo da máxima (100%). Para alcançar essa marca, a metodologia resumiu-se a seguir o seguinte roteiro: a) leitura e releitura, pelos especialistas do Direito, dos processos relacionados a cada tema (dados fornecidos pelo STF); b) extração das palavras-chave mais importantes para identificação de cada tema; c) revisão das palavras-chave pelos pares especialistas; d) apresentação do tagging para a equipe de TI; e) comparação dos dados extraídos de forma autônoma pela equipe de TI e obtidos por meio de *machine learning*, com os dados extraídos pela equipe do Direito.

Na Figura 1, a seguir, resumem-se as fases do Projeto e o conhecimento adquirido em cada etapa:

Figura 1. Fases do Projeto e o conhecimento adquirido em cada etapa



Fonte: Elaborado pelos autores.

Metodologia de extração de dados: o exemplo dos temas de repercussão geral mais relevantes e o que eles informam

Projetos em aprendizado de máquina lidam com grandes quantidades de dados e com padrões de acentuada complexidade. Quando projetos como o Victor são desenvolvidos na esfera pública, a transparência e o respeito aos valores democráticos são importantes elementos, que devem estar incluídos, inclusive, no desenho do plano executado, de forma a garantir *accountability* do cidadão beneficiário (FARRANHA, 2020).

Todos os sistemas baseados em Inteligência Artificial, e em particular os que são implantados pelo Poder Público, precisam ser seguros e responsáveis. Sua utilização deve estar sempre sujeita a controles e à avaliação contínua de riscos. Os responsáveis por seu desenvolvimento e operação devem obedecer às regulamentações jurídicas que estabeleçam com clareza a forma como tais sistemas irão impactar a sociedade. Na mesma direção, faz-se necessária a adoção de protocolos que assegurem a participação da sociedade na regulamentação da garantia da liberdade de expressão e comunicação (MAIA FILHO; JUNQUILHO, 2018, p. 225).

A oportunidade de lidar com dados no trabalho com a IA traz resultados que vão muito além da sua utilização para *machine learning* e *deep learning*, já que os dados são significativos quando adequadamente interpretados (JORDAN; MITCHELL, 2015). Como observam Giuffrida, Lederer e Vermeys (2018, p. 753), o aprendizado de máquina depende de dados: “[...] quanto mais dados puder acessar, melhor poderá ‘aprender’”. Isso é, o aprendizado de máquina usa computadores para executar modelos preditivos, que aprendem com os dados existentes para prever futuros comportamentos, resultados e tendências.

Para além da importância para projetos que utilizam a IA, saber o que os dados nos informam é fator relevante para que as partes interessadas possam ter confiança nas decisões tomadas e apresentadas como resultado desses projetos. É dizer que, nas ações implementadas em IA, é fundamental compartilhar resultados trazidos pelos dados, para o controle crítico democrático.

Por meio da análise dos dados feita por especialistas da área, foi possível identificar, por exemplo, quais foram, nos dois últimos anos, os temas mais frequentes que chegaram ao STF. A interpretação desses dados trouxe também informações relevantes sobre quais são os maiores litigantes que chegam ao STF e quais os temas de RG mais frequentes.

O Quadro 1, a seguir, apresenta, por ordem de quantidade de processos relacionados, os temas de RG mais recorrentes nos anos de 2017 a 2019 no STF:

Quadro 1. Temas de Repercussão Geral por ordem, quantidade de processos relacionados e o respectivo conteúdo, coletados entre 01/04/2017 a 30/04/2019.

Tema	Qtd. de processos	Conteúdo
660	686	Violação dos princípios do contraditório e da ampla defesa, quando o julgamento da causa depender de prévia análise da adequada aplicação das normas infraconstitucionais. Extensão do entendimento ao princípio do devido processo legal e aos limites da coisa julgada.
895	247	Ofensa ao princípio da inafastabilidade de jurisdição na hipótese em que há óbice processual intransponível ao julgamento de mérito.
810	282	Validade da correção monetária e dos juros moratórios incidentes sobre as condenações impostas à Fazenda Pública, conforme previstos no art. 1º-F da Lei 9.494/1997, com a redação dada pela Lei 11.960/2009.
852	131	Avaliação judicial de critérios para a caracterização de trabalho especial, para fins de reconhecimento de aposentadoria especial ou de conversão de tempo de serviço, nos termos dos artigos 57 e 58 da Lei 8.213/1991.
800	154	Presunção relativa de inexistência de repercussão geral dos recursos extraordinários interpostos nas causas processadas nos Juizados Especiais Cíveis da Lei 9.099/1995. Redação original: Viabilidade de recurso extraordinário contra acórdão proferido por Juizado Especial Cível da Lei 9.099/1995 em matéria de responsabilidade pelo adimplemento de obrigação assumida em contrato de direito privado.
424	114	Indeferimento de produção de provas no âmbito de processo judicial.
339	871	Obrigatoriedade de fundamentação das decisões judiciais.

766	808	Verificação dos requisitos legais necessários para concessão de benefício previdenciário.
655	792	Modificação do valor fixado a título de indenização por danos morais.
773	762	Eficácia temporal de sentença transitada em julgado fundada em norma supervenientemente declarada inconstitucional pelo Supremo Tribunal Federal em sede de controle concentrado.
634	710	Isonomia de gênero quanto ao critério de expectativa de vida adotado no cálculo do fator previdenciário.
951	486	Direito dos servidores federais às diferenças relacionadas ao reajuste de 47,11% sobre a parcela denominada adiantamento do PCCS (adiantamento pecuniário), após a mudança para o regime estatutário.
181	452	Pressupostos de admissibilidade de recursos de competência de outros Tribunais.
589	358	Revisão de renda mensal de benefício previdenciário mediante aplicação dos mesmos índices utilizados para reajuste do teto do salário-de-contribuição, relativamente aos meses de junho de 1999 e maio de 2004.
548	330	Dever estatal de assegurar o atendimento em creche e pré-escola às crianças de zero a cinco anos de idade.
188	310	Declaração de hipossuficiência para obtenção de gratuidade de justiça.
232	299	Indenização por danos morais decorrentes de inscrição indevida em cadastro de inadimplentes.
5	283	Compensação da diferença de 11,98%, resultante da conversão em URV dos valores em cruzeiros reais, com o reajuste ocorrido na data-base subsequente.
597	272	Imposição ao INSS, nos processos em que figure como parte ré, do ônus de apresentar cálculos de liquidação do seu próprio débito.
6	181	Dever do Estado de fornecer medicamento de alto custo a portador de doença grave que não possui condições financeiras para comprá-lo.
313	125	Aplicação do prazo decadencial previsto na Medida Provisória nº 1.523/97 a benefícios concedidos antes da sua edição.
163	101	Contribuição previdenciária sobre o terço constitucional de férias, a gratificação natalina, os serviços extraordinários, o adicional noturno e o adicional de insalubridade.
793	77	Responsabilidade solidária dos entes federados pelo dever de prestar assistência à saúde.
351	63	Extensão a inativos e pensionistas da Gratificação de Desempenho do Plano Geral de Cargos do Poder Executivo – GDPGPE.
33	61	Relevância e urgência da Medida Provisória nº 2.170-36/2001, que disciplina a capitalização de juros com periodicidade inferior a um ano nas operações realizadas pelas instituições integrantes do Sistema Financeiro Nacional.
555	51	Fornecimento de Equipamento de Proteção Individual – EPI – como fator de descaracterização do tempo de serviço especial.
409	48	Extensão, em relação aos servidores inativos, dos critérios de cálculo da GDPST estabelecidos para os servidores em atividade.

406	46	Critérios para o cálculo da renda mensal inicial de benefício previdenciário.
Total de processos utilizados: 22.600		

Fonte: Elaborado pelos autores.

Neste ponto, é válido perguntar: o que esses dados nos apontam? A IA é capaz de encontrar estruturas e regularidades nos dados, mediante operação que faz com que o algoritmo se torne um classificador. Ao classificar os dados em grupos, torna-se possível interpretá-los de modo mais consistente, compreendendo o conteúdo e o alcance que possuem.

No Quadro 2, foram agrupados os temas, de acordo com a separação feita na metodologia de desenvolvimento do Projeto, ou seja, por assunto:

Quadro 2. Classificação dos processos em classes de temas mais frequentes.

Área	Temas
Juizados	800
Processuais	660, 895, 424, 339, 773, 181, 188
Fazenda Pública	810, 548, 005, (006, 793 - Saúde)
Previdenciário	852, 634, 951, 589, 597, 313, 163, 351, 555, 409, 406
Instituição bancária	033
Danos Morais	655, 232

Fonte: Elaborado pelos autores.

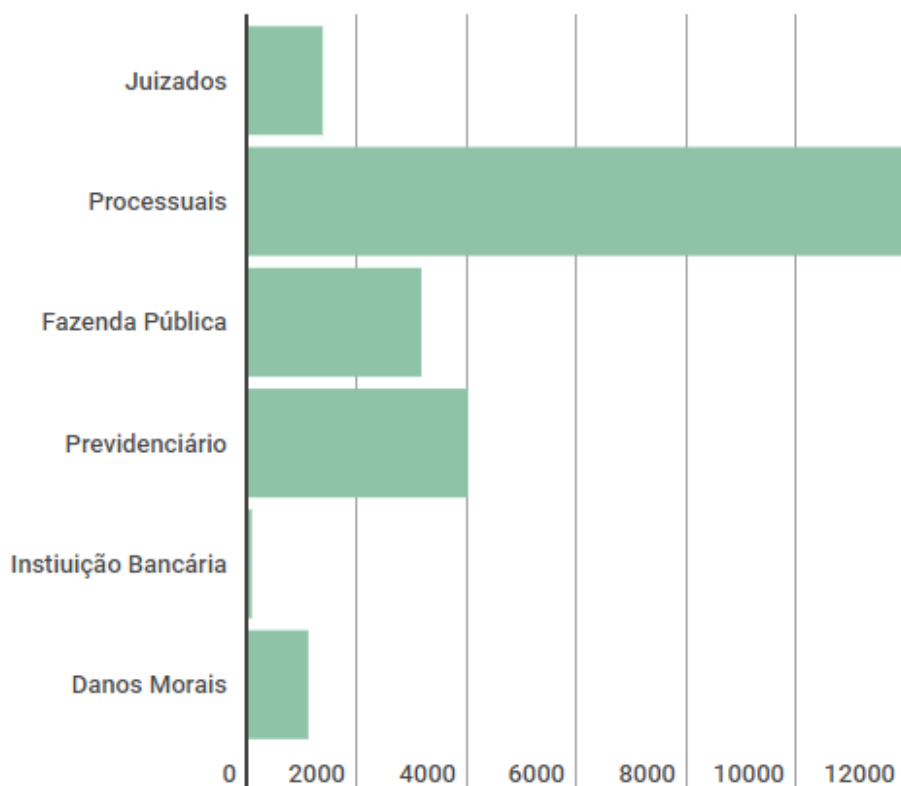
Pela classificação em áreas, observa-se que cinco dos 27 temas mais recorrentes no controle de constitucionalidade difuso têm a Fazenda Pública como parte do processo. Além disso, 12 desses 27 temas tangenciam questões previdenciárias. Essa constatação indica que, dos inúmeros processos que chegam ao STF, a maioria tem como litigante o Estado e diz respeito a discussões de direito previdenciário.

Apesar de haver apenas um tema relativo a instituições bancárias, tal fato revela que essas instituições aparecem com frequência no Poder Judiciário, alcançando as últimas vias recursais.

As questões processuais são frequentes (sete dos 27 temas), o que demonstra que as partes têm questionado o exercício da jurisdição e o correto desenvolvimento formal do processo, buscando resguardar a garantia do contraditório e da ampla defesa. O fato de um dos temas ser relativo a Juizados Especiais também denota que processos que questionam a aplicação desse procedimento especial contribuem, da mesma forma, para o alto número de processos recebidos pelo STF.

Também é possível analisar o número de processos relacionados por tema, o que leva a conclusões interessantes, como se pode ver no Gráfico 1, a seguir.

Gráfico: Temas de RG por número de processos relacionados



Fonte: Elaborado pelos autores.

Os processos advindos de Juizados totalizaram 1.354; os que abordaram questões processuais foram 12.166; 3.153 processos envolveram a Fazenda Pública em juízo; 4.006 processos abordaram temas previdenciários; 61 processos eram relativos a instituições bancárias e 1.091, dos mais frequentes, tangenciavam o tema de danos morais.

As questões que mais apareceram nos processos analisados discutiam, portanto, o respeito ao devido processo legal, seguidos por aqueles cujo conteúdo era previdenciário. Em resumo, os dados trouxeram informações que indicam: a) as áreas que demandam maior investimento em políticas públicas; b) quais os maiores litigantes da Justiça; c) as questões processuais que têm sido mais discutidas na Corte Constitucional etc. Essas são algumas demonstrações do que pode ser obtido como informação relevante em projetos na área de IA.

Sabe-se que, do aprendizado de máquina, pode-se extrair, de forma automática, um grande volume de dados complexos, assim como indicadores com elevado grau de precisão. O Projeto Victor, nesse sentido, contribui para aperfeiçoar a execução de tarefas rotineiras dos servidores e para tomadas de decisão mais apropriadas, reduzindo drasticamente o tempo gasto em buscas jurisprudenciais e em documentos internos ao processo. O processo decisório constitucional difuso ganha apoio do AM, o qual traduz ganhos em acerto, exatidão e rigor na correção dos julgamentos, contribuindo para a redução do tempo de duração desses processos.

Como se abordou inicialmente, o *design* para execução de projetos de IA que reflitam bons resultados para a sociedade envolve diversos fatores, entre os quais a transparência, a explicação dos objetivos pretendidos e alcançados e a explanação dos dados utilizados (FLORIDI *et al.*, 2020). Por isso, há importância nas análises e em estudos como este, que apresentem à sociedade resultados de pesquisas em IA, como a que foi realizada no âmbito do Projeto Victor.

Considerações Finais

A Inteligência Artificial aplicada ao campo do Direito, como visto, apresenta boas perspectivas de incremento da funcionalidade e do aprimoramento da prestação jurisdicional e do acesso à Justiça. Por meio do treinamento e da inserção de novos dados, que o STF deverá continuar em etapas seguintes, o Projeto Victor facilitará sobremaneira o trabalho dos ministros e servidores do STF. Permitirá a execução de tarefas e a operação de sistemas com uma precisão que, frente ao exacerbado volume de processos existentes, é hoje impossível de ser alcançada por meios meramente manuais.

Em uma Corte que lida com elevados índices de litigiosidade como é o STF, a utilização de métodos de *machine learning* pode trazer maior velocidade no tratamento e processamento de dados, maior grau de precisão na análise das informações coletadas e um processo de tomada de decisões mais célere e eficaz. Esse processo, inevitavelmente, levará à melhoria dos serviços jurisdicionais prestados aos cidadãos.

Os resultados até aqui alcançados pelo Projeto Victor permitem concluir que o investimento em novas tecnologias de computação cognitiva oferece condições para que o Judiciário experimente uma desburocratização de seus serviços e o incremento da produtividade dos juízes e tribunais. Como evidenciado, a automatização do fluxo de informações permite a racionalização do tempo gasto na classificação, organização e acompanhamento de processos e uma qualificada coleta de dados.

O Projeto Victor conseguiu revelar os maiores litigantes da Justiça e os processos repetitivos, assim como as questões processuais mais recorrentes nas ações que chegam ao STF. Além disso, trouxe informações qualificadas e seguras, possibilitou a indexação de conteúdos, a detecção de erros e a canalização de esforços de seu corpo técnico para funções mais especializadas e criativas. Tais elementos são decisivos para que as decisões judiciais sejam tomadas de maneira mais consistente e para que se coloquem freios à avassaladora onda de processos que assoberbam a Corte Suprema, muitos deles sem qualquer significação constitucional.

As melhorias na qualidade dos serviços públicos prestados – aí incluída a prestação jurisdicional – necessitam de ferramentas que permitam planejamentos mais eficazes e uma melhor governança. A AI4SG, nesse contexto, pode oferecer mecanismos para que o Poder Público invista no aperfeiçoamento dos sistemas de gestão, combatendo a burocratização e a baixa produtividade na Administração Pública.

A economia dos algoritmos já é uma realidade em inúmeros setores e, nesse sentido, é fundamental que os países periféricos, em particular, não se atrasem no domínio da IA, como já ocorreu anteriormente com relação a outras tecnologias. Ao mesmo tempo, é necessário estabelecer mecanismos de regulação jurídica que garantam sua utilização de forma ética e responsável. De modo particular, as instituições públicas devem trabalhar para que os benefícios da IA estejam voltados à melhoria do acesso à informação, à promoção de melhores serviços governamentais e à garantia da participação social. Por isso, é importante que se tenha em mente a ideia de AI4SG, isso é, uma IA que se preocupe com o *design* ético e que seja planejada para atender às especificidades da comunidade à qual se aplica, que seja transparente e apresente seus resultados à população.

A implantação do Projeto Victor visou, desde seu início, à otimização dos julgamentos e à eficiência da prestação jurisdicional no âmbito do STF. A captura e a análise de dados que o Projeto realizou contribuirão consideravelmente no combate à cultura da hiperlitigância e na identificação de irrelevância processual, particularmente em temas de repercussão geral. Por isso, salienta-se, mais uma vez, a importância de pesquisas como esta, que mostrem e compartilhem seus resultados com o mundo jurídico e com a sociedade em geral, destinatária final que é das políticas públicas e dos avanços tecnológicos que a inteligência humana tem sido capaz de conceber.

Referência

CUEVAS, Ricardo Villas Bôas. Inteligência Artificial no Judiciário. In: **Inteligência Artificial e Direito Processual: Os Impactos da Virada Tecnológica no Direito Processual**. Coordenação de Dierle Nunes, Paulo Henrique dos Santos Lucon e Erick Navarro Wolkart. Salvador: Juspodvim, 2021, p. 79-91.

DOMINGOS, Pedro. A Few Useful Things to Know About Machine Learning. In: **Communications of the ACM**, v. 55, 2012, p. 78-87.

FARRANHA, Ana Cláudia. Serviços Públicos e a Estratégia Brasileira de Inteligência Artificial: primeiras abordagens. In: BRAVO, Álvaro Sánchez. **Derecho, Inteligencia Artificial y Nuevos entornos digitales**. Espanha: Álvaro Sánchez Bravo, 2020.

FEDERAL CONSTITUTIONAL COURT. **Anual Statistics 2016**. Karlsruhe, 2017. Disponível em: https://www.bundesverfassungsgericht.de/SharedDocs/Downloads/EN/Statistik/statistics_2016.pdf;jsessionid=7AFD8EB2C71FBF16ADCD79149E4074CD.2_cid393?__blob=publicationFile&v=2. Acesso em: 03 mar. 2021.

FLORIDI, L. *et al.* How to Design AI for Social Good: Seven Essential Factors. In: **Science and Engineering Ethics**. Springer Netherlands, 2020, p. 1771-1796.

GHANI, Rayid. **You Say You Want Transparency and Interpretability?** Rayid Ghani (blog) 2016. Disponível em: <http://www.rayidghani.com/you-say-you-want-transparency-andinterpretability>. Acesso em: 03 fev. 2020.

GIUFFRIDA, I. A legal perspective on the trials and tribulations of Ali: how artificial intelligence, the internet of things, smart contracts, and other technologies will affect the law. **Case Western Reserve Law Review**, 68(3), 2018, p. 747-782.

HAGER, G. D., DROBNIS, A., FANG, F., GHANI, R., Greenwald, A., Lyons, T., & Parkes, D. C. *et al.* **Artificial intelligence for social good**, 2017.

HULDTGREN, Alina. Design for Values in ICT. In: HOVEN, Jeroen van den; VERMAAS, Pieter E.; POEL, Ibo van de. **Handbook of Ethics, Values and Technological design: sources, theory, values and application domains**. Nova Iorque: Springer, 2015, p. 739-768.

JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. **Science**, v. 349, n. 6245, p. 255-260, 2015.

KROES, Peter; POEL, Ibo van de. Design for Values and the Definition, Specification, and Operationalization of Values. In: HOVEN, Jeroen van den; VERMAAS, Pieter E.; POEL, Ibo van de. **Handbook of Ethics, Values and Technological design: sources, theory, values and application domains**. Nova Iorque: Springer, 2015.

LAVILLE, Christian; DIONNE, Jean. **A construção do saber: Manual de metodologia da pesquisa em ciências humanas**. Porto Alegre: Editora Artes Médicas, 1999.

LEE, Kai-fu. **AI Super-Powers: China, Silicon Valley, and the New World Order**. Nova Iorque: Houghton Mifflin, 2018.

LIMA, Flávia Danielle Santiago; ANDRADE, Louise Dantas de; OLIVEIRA, Tassiana Moura de. As (in) alterações do novo código de processo civil na repercussão geral: apontamentos sobre a atuação do STF. **Revista CEJ**, Brasília, Ano XIX, n. 67, p. 78-84, set. 2015.

LOURIDAS, P., & EBERT, C. Machine Learning. **IEEE Software**, 2016, p. 110-115.

MAGRANI, E. **Entre dados e robôs: ética e privacidade na era da hiperconectividade**. Porto Alegre: Arquipélago Editorial, 2019.

MAIA FILHO, Mamede Said; JUNQUILHO, Tainá Aguiar. Projeto Victor: Perspectivas de aplicação da inteligência artificial ao Direito. **Revista de Direitos e Garantias Fundamentais**. Vitória, v. 19, n. 3, 2018, p. 219-238.

MEADOR, Daniel John; MITCHELL, Gregory. **American Courts**. 3. ed. Saint Paul: West Academic Publishing, 2009.

PEREZ, J. A., DELIGIANNI, F., RAVI, D., & YANG, G. Artificial Intelligence and Robotics. **EUK-RAS White Paper Series on Robotics and Autonomous Systems (RAS)**, 2017.

RUSSELL, Stuart. **Human Compatible: Artificial Intelligence and the Problem of Control**. Estados Unidos: Viking, 2019.

SILVA, Nilton Correia da. Notas iniciais sobre a evolução dos algoritmos do Victor: o primeiro projeto em inteligência artificial em supremas cortes do mundo. In: FERNANDES, Ricardo Vieira de Carvalho; CARVALHO, Angelo Gamba Prata de (Coord.). **Tecnologia jurídica & direito digital: II Congresso Internacional de Direito, Governo e Tecnologia - 2018**. Belo Horizonte: Fórum, 2018.

SCHWAB, Klaus. **A quarta revolução industrial**. Edipro, 2019.

SUPREMO TRIBUNAL FEDERAL. **Relatório de Atividades 2020**. Brasília: 2021. Disponível em: <http://www.stf.jus.br/arquivo/cms/publicacaoCatalogoProdutoConteudoTextual/anexo/RelatorioAtividadesSTF2020.pdf>. Acesso em: 7 mai. 2021.

TRIVIÑOS, A. N. S. (2013). Introdução à Pesquisa em Ciências Sociais: a Pesquisa Qualitativa em Educação. In: TRIVIÑOS, Augusto Nibaldo Silva. **O Positivismo, A Fenomenologia, O Marxismo**. (Vol. 1, n. 20). Disponível em: <https://doi.org/10.33081/formação.v1i20.2335>. Acesso em: 7 jun. 2020.

VIEHWEG, Theodor. **Tópica e Jurisprudência**. Trad. Tércio Sampaio Ferraz Jr. Brasília: Departamento de Imprensa Nacional, 1979.

Recebido em 24 de maio de 2021

Aceito em 16 de junho de 2021