



Programa de Mestrado Profissional em Direito, Regulação e Políticas Públicas  
Faculdade de Direito  
Universidade de Brasília

GUILHERME SILVA FIGUEIREDO

**PROJETO ATHOS: Um Estudo de Caso sobre a inserção do  
Superior Tribunal de Justiça na Era da Inteligência Artificial.**

Orientador:

Prof. Dr. Alexandre Araújo Costa

**Brasília**

**2022**

**GUILHERME SILVA FIGUEIREDO**

**PROJETO ATHOS: Um Estudo de Caso sobre a inserção do Superior Tribunal de Justiça na Era da Inteligência Artificial.**

Trabalho de conclusão de curso apresentado à Coordenação do Curso de Direito da Universidade de Brasília- UnB, como requisito parcial para obtenção do grau de Mestrado Profissional em Direito, Regulação e Políticas Públicas. Orientador: Prof. Dr. Alexandre Araújo Costa.

**Brasília**

**2022**

Figueiredo, Guilherme Silva.

Título: PROJETO ATHOS: Um Estudo de Caso sobre a inserção do Superior Tribunal de Justiça na Era da Inteligência Artificial – Brasília, 2022.

Trabalho de conclusão de curso apresentado à Coordenação do Curso de Direito da Universidade de Brasília- UnB, como requisito parcial para obtenção do grau de Mestrado Profissional em Direito, Regulação e Políticas Públicas. Orientador: Prof. Dr. Alexandre Araújo Costa.

1. Inteligência Artificial. 2. Direito. 3. Recursos Repetitivos. 4. Projeto Athos.

CDU

**GUILHERME SILVA FIGUEIREDO**

**PROJETO ATHOS: Um Estudo de Caso sobre a inserção do Superior Tribunal de Justiça na Era da Inteligência Artificial.**

Trabalho de conclusão de curso apresentado à Coordenação do Curso de Direito da Universidade de Brasília- UnB, como requisito parcial para obtenção do grau de Mestrado Profissional em Direito, Regulação e Políticas Públicas. Orientador: Prof. Dr. Alexandre Araújo Costa.

Brasília, \_\_\_\_\_ de \_\_\_\_\_ de 2022.

Banca Examinadora

---

ALEXANDRE ARAÚJO COSTA  
Doutor  
UnB- Universidade de Brasília

---

Doutor  
UnB- Universidade de Brasília

---

Doutor  
UnB- Universidade de Brasília

NOTA: \_\_\_\_\_

Para Tatiany, Amanda e Rafael.  
Vocês são o meu destino.

## AGRADECIMENTOS

Aos meus pais, Raimundo (*in memoriam*) e Luiza, por iniciarem, através de muito trabalho e dedicação, esta conquista. Muito obrigado!

Aos colegas do Núcleo de Gerenciamento de Precedentes e de Ações Coletivas (NUGEPNAC), da Secretaria Judiciária (SJD) e da Assessoria de Inteligência Artificial (AIA), do Superior Tribunal de Justiça, pelos préstimos em disponibilizar os dados necessários à realização desse trabalho.

## RESUMO

O avanço digital e os novos desafios impostos pelas tecnologias 4.0 são inegáveis. Para tanto, necessitamos analisar, de forma profunda e com olhar crítico, as ferramentas que se apresentam como solução para os problemas, sejam de segurança pública ou demandas judiciais. O uso de sistemas de Inteligência Artificial é crescente nas mais diversas áreas, em razão do aumento da eficiência e da precisão dos serviços por eles proporcionado. No Direito, ocorre também esse fenômeno. Ademais, no Judiciário, foram iniciados diversos estudos nesse sentido numa tentativa de aumentar a efetividade, a eficiência e a produtividade dos serviços prestados à sociedade. Dados do Conselho Nacional de Justiça informam que a taxa de congestionamento do Poder Judiciário, que mede o percentual de processos que ficam represados sem solução ao final de cada ano, segue em patamares insustentáveis. Com vistas a mitigar a vazão de processos ingressados anualmente, o Superior Tribunal de Justiça implementou sua solução de Inteligência Artificial: o projeto Athos. O presente estudo objetivou realizar um Estudo de Caso sobre esta ferramenta. Apresentada em 2019, essa solução de inteligência artificial, possui sua principal vertente utilizada no Núcleo de Gerenciamento de Precedentes e de Ações Coletivas (NUGEPNAC), e consiste no monitoramento e agrupamento de processos, com vistas a criação de novos temas repetitivos daquela Corte. Fundamentada em vasta pesquisa bibliográfica, o estudo iniciou com a coleta de dados na base de temas repetitivos do Tribunal, e em dados de indicadores estratégicos disponibilizados pela equipe do NUGEPNAC e da Assessoria de Inteligência Artificial (AIA). Posteriormente, efetuou-se a análise dos dados com o apoio de ferramentas de *Business Intelligence* (BI), e a apresentação dos respectivos resultados. Objetiva-se que, a partir desses resultados, a pesquisa possa servir de base para novos estudos.

**Palavras-chave:** Superior Tribunal de Justiça. Inteligência Artificial. Recursos Repetitivos. Projeto Athos.

## ABSTRACT

It is impossible to deny the digital advancement and the new challenges that 4.0 technologies impose on us. Therefore, we need to analyze, in depth and with a critical eye, the tools that are presented as a solution to problems, whether in public security or legal claims. The use of Artificial Intelligence systems is growing in the most diverse areas, due to the increase in efficiency and precision of the services they provide. In Law, this phenomenon also occurs. In the Judiciary, several studies were initiated in this regard in an attempt to increase the effectiveness, efficiency and productivity of services provided to society. Data from the National Council of Justice inform that the Judiciary Branch's congestion rate, which measures the percentage of cases that remain unresolved at the end of each year, remains at unsustainable levels. In order to mitigate the flow of cases filed annually, the Superior Court of Justice implemented its Artificial Intelligence solution: the Athos project. The present study aimed to carry out a case study on this tool. Introduced in 2019, this artificial intelligence solution, has its main aspect used in the Center for Management of Precedents and Collective Actions - NUGEPNAC, and consists of monitoring and grouping of processes, with a view to creating new repetitive themes of that Court. Based on extensive bibliographical research, the study began with the collection of data based on repetitive themes of the Court, and on data from strategic indicators made available by the NUGEPNAC team and the Assessoria de Inteligência Artificial (AIA). Subsequently, data analysis was carried out with the support of Business Intelligence (BI) tools, and the presentation of the respective results. The aim is that, based on its results, the research can serve as a basis for further studies.

**Keywords:** Superior Court of Justice. Artificial intelligence. Repetitive Features. Athos Project.



## LISTAS DE FIGURAS

Figura 1- Jurimetria no contexto do Direito tecnológico

Figura 2- Fluxo CRISP-DM.

Figura 3- Tela do filtro de busca por processos semelhantes do sistema Athos.

Figura 4- Tela consultar processos semelhantes do sistema Athos.

Figura 5- Tela do sistema Athos com o resultado busca por similaridade.

Figura 6- Tela da busca textual do sistema Athos.

Figura 7- Tela busca por semelhança do sistema Athos.

Figura 8- Tela busca monitoramento de grupos do sistema Athos.

## **LISTA DE TABELAS**

Tabela 1- Características do Algoritmo.

Tabela 2- Classificação dos Algoritmos.

Tabela 3- Quantidade de controvérsias criadas entre os anos de 2020 e 2021 pelo NUGEPNAC.

Tabela 4- Redução de Recursos Especiais partir do ano de 2016

## LISTA DE GRÁFICOS

- Gráfico 1- Transformação das carreiras.
- Gráfico 2- Processos julgados a partir de temas do rito dos repetitivos
- Gráfico 3- Série histórica processos julgados RR-STJ
- Gráfico 4-Tribunais que enviam processos com temas definidos
- Gráfico 5- Processos ingressados anualmente no STJ
- Gráfico 6- Atividade jurisdicional do STJ entre janeiro e julho de 2021
- Gráfico 7- Distribuição de incidentes no poder judiciário
- Gráfico 8- Número de incidentes ingressados anualmente/número de incidentes julgados.
- Gráfico 9- Número de incidentes ingressados anualmente/número de processos sobrestados.
- Gráfico 10- Ramos do Direito dos temas cadastrados na base de repetitivos do STJ.
- Gráfico 11- Contagem de Temas/Quantidade de processos suspensos no STJ.
- Gráfico 12- Série histórica da quantidade de temas cadastrados na base de dados de precedentes qualificados do STJ.
- Gráfico 13- Número de processos sobrestados em todas instancias do judiciário.
- Gráfico 14- Tipos de RRC cadastrados na base de dados do STJ.
- Gráfico 15- Quantidade de grupos monitorados pelo NUGEPNAC.
- Gráfico 16- Ramos do Direito nos processos estudados pelo NUGEPNAC.
- Gráfico 17- Status dos grupos monitorados nos anos de 2020 e 2021 pelo NUGEPNAC.
- Gráfico 18- Quantidade de Temas cancelados cadastrados na base de dados do STJ.
- Gráfico 19- Quantidade de Grupos/Controvérsias/Temas criados pelo sistema Athos entre os anos de 2020 e 2021.
- Gráfico 20- Quantidade de funcionalidades utilizadas no sistema Athos entre os anos de 2020 e 2021.
- Gráfico 21- Funcionalidades Rodar Acervo/Agrupamentos utilizadas no sistema Athos entre os anos de 2020 e 2021.
- Gráfico 22- Quantidade de grupos criados pelo Athos e NUGEPNAC entre os anos de 2020 e 2021.
- Gráfico 23- Série histórica da quantidade de processos recebidos no STJ.
- Gráfico 24- Recorte dos principais recursos recebidos no STJ, a partir do ano de 2016

## LISTA DE ABREVEATURA E SIGLAS

<b>AIA</b>	Assessoria de Inteligência Artificial
<b>AM</b>	Aprendizado de máquina
<b>AREsp</b>	Agravos em Recurso Especial
<b>BNPR</b> Obrigatórios	Banco Nacional de Dados de Demandas Repetitivas e Precedentes
<b>CNJ</b>	Conselho Nacional de Justiça
<b>COGEPAC</b>	Comissão Gestora de Precedentes e de Ações Coletivas
<b>CPC</b>	Código de Processo Civil
<b>CRISP-DM</b>	Cross Industry Standard Process for Data Mining
<b>HC</b>	Habeas Corpus
<b>IA</b>	Inteligência Artificial
<b>IN</b>	Instrução Normativa
<b>IRDR</b>	Resolução de Demandas Repetitivas
<b>MP</b>	Ministério Público
<b>NARER</b>	Núcleo de Admissibilidade e Recursos Repetitivos
<b>NUGEPNAC</b>	Núcleo de Gerenciamento de Precedentes e de Ações Coletivas
<b>OiT</b>	Internet das Coisas
<b>PIB</b>	Produto Interno Bruto
<b>PLN</b>	Processamento de Linguagem Natural
<b>REsp</b>	Recurso Especial
<b>RISTJ</b>	Regimento Interno do STJ
<b>RR</b>	Recursos Repetitivos
<b>RRC</b>	Recurso Representativo de Controvérsia
<b>SJD</b>	Secretaria Judiciária
<b>SJR</b>	Secretaria de Jurisprudência do STJ
<b>STF</b>	Supremo Tribunal Federal
<b>STI</b>	Secretaria de Tecnologia da Informação e Comunicação do STJ
<b>STJ</b>	Superior Tribunal de Justiça
<b>TFR</b>	Tribunal Federal de Recursos
<b>TUA</b>	Tabela Unificada de Assuntos

# Sumário

Introdução.....	13
1. Inteligência artificial .....	17
1.1. A quarta revolução industrial .....	17
1.2. Algoritmos .....	25
1.3. Modelos de inteligência artificial .....	30
1.3.1. Machine learning .....	30
1.3.2. REDES NEURAIS E DEEP LEARNING .....	35
1.3.3. PROCESSAMENTO DE LINGUAGEM NATURAL (PLN) .....	38
1.3.4. CORPUS .....	39
1.3.5. Paragraph Vector .....	40
2. Inteligência artificial no Poder Judiciário .....	41
2.1. Inteligência Artificial e decisões judiciais.....	41
2.2. Inteligência artificial no Poder Judiciário .....	45
3. Impactos do Projeto Athos no processo decisório do STJ .....	50
3.1. Introdução o STJ inserido na era da inteligência artificial.....	50
3.2. Modelagem dos dados .....	53
3.2.1. Conversão para minúsculas .....	55
3.2.2. Remoção de <i>StopWords</i> .....	55
3.2.3. Remoção de termos entre parênteses:.....	55
3.2.4. Remoção de Não-letas:.....	56
3.2.5. Singularização de Verbetes:.....	56
3.2.6. Conversão de caracteres especiais:.....	56
3.3. O sistema Athos e o rito dos recursos repetitivos .....	59
4.0 As Demandas do Poder Judiciário e Os Resultados do Sistema Athos .....	66
4.1 A crise de demandas no poder judiciário e os recursos repetitivos.....	66
4.2 O sistema Athos no apoio a criação de novos temas repetitivos .....	77
5. Conclusão.....	103
6. Referências.....	106

# Introdução

As tentativas de compreender e modelar as habilidades cognitivas humanas datam de séculos, havendo registros históricos de numerosos tentativas de se entender os princípios do raciocínio e da memorização, tanto do ponto de vista filosófico, quanto da fisiologia e da biologia do cérebro humano. (SILVA, 2020). Já as tentativas de projetar ferramentas capazes de processar comandos (os computadores) se desenvolveram a partir do século XIX, com Charles Babbage, assim como a produção de algoritmos voltados a serem implementados por essas máquinas, sendo que o um algoritmo publicado por Ada Lovelace é considerado o primeiro “software”.

Embora a ideia de que seria viável produzir máquinas inteligentes esteja ligada a essas tentativas, a expressão “inteligência artificial” foi cunhada apenas em 1956, por John McCarthy, na primeira conferência organizada por ele e Marvin Minsky sobre o assunto (*Dartmouth Summer Research Project on Artificial Intelligence - DSRPAI*)<sup>1</sup>. Desde então, a ideia de uma inteligência artificial e seus impactos moveu uma série de cientistas e artistas (como Isaac Asimov), muito antes que fossem construídos os computadores que seriam capazes de implementar esse tipo de estratégia.

A própria discussão sobre o que seria considerado “inteligência artificial” é longa e variada, sendo que se consolidou chamar por esse nome todo sistema computacional que resolve problemas que os humanos enfrentam por meio do uso de suas habilidades cognitivas, por meio de tomadas de decisão baseadas em análises probabilísticas (MULHOLLAND, 2020). Esse é um conceito que não oferece apenas critérios para se avaliar um algoritmo e suas propriedades, mas que realiza uma espécie de analogia: toda estratégia capaz de substituir a cognição humana, em certas atividades que os humanos realizam por meio de sua inteligência, pode ser considerada inteligência artificial (IA).

Isso não significa que os computadores sejam efetivamente inteligentes, nem conscientes, nem que eles realizem algo próximo da cognição

---

<sup>1</sup> MOOR, J. The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years: AAAI AI Magazine, v. 27, n. 4, 2006.

humana. O que interessa é o resultado: toda estratégia computacional capaz de substituir a inteligência humana por outras formas de abordagem (normalmente por algoritmos baseados em análises probabilísticas) pode ser considerada como uma estratégia de IA.

Todavia, devemos tomar aqui alguns cuidados, pois um conceito muito abrangente de IA terminaria por abarcar uma série de atividades que não são ligadas a esse conceito, mas que são feitas pelos humanos por meio de processos que envolvem cognição, como a busca, a leitura e o armazenamento de dados.

O uso intensivo de computadores permite que a IA opere mediante o rápido processamento de uma vasta gama de informações. Os dados, uma vez analisados, levam à possibilidade de tomadas de decisão, tanto por humanos, quanto pelas próprias “máquinas”. O objetivo do desenvolvimento de IA é a aceleração de processos de aprendizado e a otimização de seus resultados, visando uma maior eficiência e reduzindo o tempo de análise de dados necessário para a tomada de decisão (MULHOLLAND, 2020).

Dados referentes à evolução temporal da área de IA mostram um avanço significativo a partir do final dos anos 1990. Verificam-se aumentos notáveis na produção científica (o crescimento de publicações anuais de artigos em IA supera muito a quantidade geral publicada na área de Ciência da Computação (CC) entre 1996 e 2017, oferta de vagas de trabalho, abertura de novas empresas (entre 2015 e 2018 houve um aumento de startups em IA na ordem de 113%, bem como em investimentos financeiros (venture capital). (SILVA, 2020) Este notório desenvolvimento foi impulsionado por alguns principais fatores, como a própria evolução metodológica da área – modelos matemáticos que permitem a mimetização de funções humanas de reconhecimento visual, de voz e de textos, bem como da função humana de correlacionar informações para tomadas de decisões. (SILVA, 2020) A redução dos custos computacionais, agregada à oferta cada vez maior de dados multimídia (vídeos, imagens, áudio, texto e números) também são fatores importantes que contribuíram com o sucesso dos modelos de IA, os quais demandam muito poder de processamento e grandes quantidades de informação para serem assertivos em tarefas como: classificação de textos, de

reconhecimento de faces, de reconhecimento de voz, de predições e de agrupamentos de informações.

IA aplicada é uma realidade emergente, considerando-se especialmente a abordagem de desenvolvimento de tecnologias inteligentes por meio de aprendizado de máquina (*machine learning*) e desenvolvimento de sistemas inteligentes baseados em dados. Essa abordagem tem encontrado campo fértil em situações de tarefas que estão sendo desenvolvidas por humanos em um ambiente rico em dados específicos e confiáveis. Isso tem levado a se entender que, uma vez que se tenha um legado histórico de resolução de um tipo de problema de forma manual, esta experiência pode ser modelada e transferida para um sistema computacional que pode escalar significativamente a resolução de tal problema. (SILVA, 2020).

Vale destacar, porém, que a trajetória da IA a longo prazo dependerá do avanço da maturidade de desenvolvimento deste segmento dentro das organizações, pois, a realidade da implantação de Inteligência Artificial é muito mais limitada. Pesquisas recentes descobriram que as organizações costumam fazer experiências com Inteligência Artificial, mas lutam para tornar estas tecnologias parte de suas operações padrão. As mesmas pesquisas preveem que levará até 2025 para metade das organizações em todo o mundo alcançar o que o modelo de maturidade de Inteligência Artificial descrita como o “estágio de estabilização” de amadurecimento do uso da IA<sup>2</sup>.

No Poder Judiciário, existe uma grande expectativa de que a IA promova transformações da atividade judiciária nos próximos anos, com a criação do que já se denomina justiça digital ou justiça preditiva, o que tem elevado as reflexões acerca do impacto da revolução digital sobre o fazer jurídico. (CUEVA, 2021). Aliada a esta perspectiva, diante do crescimento exponencial das ciências cognitivas e da ciência de dados – *Data Science* –, e do enorme impacto da IA no trabalho, sobretudo em ganho de tempo, soma-se a possibilidade de, no futuro próximo, as ferramentas de IA venham a ser usadas nas tomadas de decisões judiciais, como já ocorre em alguns países,

---

<sup>2</sup> ABES- Associação Brasileira das Empresas de Software. Disponível em: <https://abessoftware.com.br/gartner-preve-que-o-mercado-global-de-inteligencia-artificial-chegara-a-us-625-bilhoes-em-2022/>. Acesso em: 05 fev. 2022



notadamente Estados Unidos e Canadá, sobretudo no campo mais sensível do Direito, que diz respeito à liberdade da pessoa, o da Justiça Criminal (SULOCKY, 2020).

Nos tribunais superiores, existem experimentos que se valem de ferramentas embrionárias de inteligência artificial, que se aproximam, ao menos, até o momento, do estágio da automação de rotinas e procedimentos, uma vez que se propõe simplesmente a classificar os recursos ingressados de acordo com listas predefinidas de categorias processuais (CUEVA, 2021). Diante desse cenário, podemos vislumbrar um judiciário equipado com ferramentas de IA capazes de fornecer suporte em suas tomadas de decisão mais críticas, ou seu uso será simplesmente restrito à automação de rotinas?

O presente estudo realizou um Estudo de Caso da ferramenta de Inteligência Artificial do Superior Tribunal de Justiça. A partir dos dados disponibilizados pelo Núcleo de Gerenciamento de Precedentes e de Ações Coletivas (NUGEPNAC) e da Assessoria de Inteligência Artificial (AIA) do Superior Tribunal de Justiça (STJ) foi possível, com auxílio de ferramentas de *Business Intelligence* (BI) tabular os resultados do Sistema Athos nos anos de 2020 e 2021, e assim posicioná-lo perante as expectativas que a IA impõe ao Direito e ao Poder Judiciário Brasileiro.

# 1. Inteligência artificial

A devida compreensão do presente trabalho exige o domínio de uma série de conceitos relativos à inteligência artificial, o que torna conveniente iniciarmos por um capítulo introdutório, que exponha essas questões e defina os conceitos que serão utilizados no restante do trabalho, para mapear as iniciativas do STJ relativas à implementação de estratégias de Inteligência Artificial aplicada à prática judiciária.

## 1.1. A quarta revolução industrial

São identificados como revolucionários aqueles acontecimentos efetivamente disruptivos da história, ou seja, aqueles que representam uma ruptura, uma descontinuação de uma determinada forma de manifestação do comportamento social. A invenção da maquinaria e, notadamente, do motor a vapor e de ferramentas sintetizadoras da técnica humana (como o tear de fição mecanizada) são identificados como acontecimentos determinantes deste processo disruptivo, alcunhado de revolução industrial. Outros acontecimentos que se sucederam também vieram a ser catalogados pelos historiadores como disruptivos. A partir daí convivemos, hoje, com consenso científico no sentido de se dividir a Revolução Industrial em fases. (RODRIGUES, 2021).

A primeira Revolução Industrial ocorreu quando a mecanização e a potência da água e do vapor multiplicaram a eficiência das tecnologias produtivas, que antes dependiam dos trabalhos humano e animal (STEIBEL; VICENTE; JESUS, 2021). A Revolução Industrial iniciada no século XVIII trouxe máquinas que substituíram o trabalho físico e ganhos de produtividade, como a máquina a vapor, força que dirigiu a inovação na indústria, na agricultura e nos transportes. Contudo, a máquina era controlada de perto pelo homem, e faltava a ela uma base de conhecimento ou inteligência<sup>3</sup>. Cabe destacar ainda que a primeira Revolução Industrial não se baseou apenas na busca permanente do homem por novas fronteiras, mas nas necessidades de defesa num contexto de

---

<sup>3</sup> MUSPRATT, Murray A. Artificial intelligence. Journal of Professional Issues in Engineering, v. 112, n. 3, p. 158-169, 1986

rivalidade entre grandes potências durante a Guerra Fria; a complexidade crescente do ambiente de formulação da decisão; e os custos crescentes de mão de obra.

Já a segunda Revolução foi marcada pela produção em massa e pelo estilo de produção marcado pela linha de montagem, facilitada pela difusão da eletricidade, que permitiu tornar mais eficiente a produção das fábricas (STEIBEL; VICENTE; JESUS, 2021). Para Teixeira e Cheliga (2020, p. 41), esta revolução é vista como apenas uma etapa da revolução industrial, pois, de um ponto de vista social e tecnológico, não houve uma clara ruptura entre as duas, sendo uma o aperfeiçoamento da outra. Os autores concluem:

O fim da segunda revolução industrial marca o início de uma nova era no cenário da indústria mundial, a era atômica. Essa era se inicia com o estudo dos átomos e tem seu ápice com as bombas nucleares que foram lançadas em Hiroshima e Nagasaki. Logicamente que os estudos em relação a física atômica não se restringem aos fins bélicos; também são importantes para o uso energético. (TEIXEIRA; CHELIGA, 2020, p. 41).

Historicamente, a terceira revolução foi marcada no período de 1945, com o fim da segunda guerra mundial, até a virada do século XXI. Teixeira e Cheliga (2020, p. 44) tratam como pontos essenciais da terceira revolução o uso de computadores e a digitalização de informações, o avanço da eletrônica, a automatização de fábricas, a corrida espacial e o avanço dos trabalhos em relação à energia nuclear. Tecnicamente, Steibel; Vicente e Jesus (2021) caracterizam a terceira revolução pelo advento dos computadores e da automação. Os autores marcam este momento como o ponto onde o poder de processamento dos computadores e a ampliação do acesso à memória foram idealmente planejados para conduzir tarefas automáticas e repetitivas, como a entrega de informação. Já em meados da década de 1980, uma quinta geração de computadores, desenvolvida com hardwares baseados em válvulas e tubos a vácuo, transistores, circuitos integrados/semicondutores e circuitos integrados bem extenso apareceu.

As primeiras explorações sobre o potencial da Inteligência Artificial datam da década de 1930, e os primeiros mapeamentos teóricos já vinham sendo realizados aproximadamente desde a década de 1950. Entretanto, até meados da década de 1980, havia carência de dados,

capacidade de processamento limitada dos recursos disponíveis e desenvolvimento precário de softwares para que a Inteligência Artificial pudesse atingir todas as suas possibilidades e potencialidades. As faces do que se considerava Inteligência Artificial eram, naquele momento, a pressão da procura e o impulso tecnológico (STEIBEL; VICENTE; JESUS, 2021).

Já a quarta Revolução Industrial envolve sistemas ciberfísicos, nos quais os humanos interagem com máquinas, e se coloca a “internet das coisas”, na qual as máquinas interagem entre si. A principal diferença entre essa atual fase e os avanços tecnológicos anteriores é a grande interação entre os mundos físico, digital e biológico. Um dos aspectos fundamentais da quarta Revolução Industrial é a ampliação das possibilidades e dos potenciais de aplicação da Inteligência Artificial (STEIBEL; VICENTE; JESUS, 2021). O ponto essencial da quarta revolução é o uso não mais centralizado do avanço das tecnologias, mas sim no avanço da comunicação e conectividade (TEIXEIRA; CHELIGA, 2021).

Três tendências são referência digitais dentro dessa revolução industrial, na categoria física podemos citar não somente os carros, mas também barcos, drones, helicópteros, tratores e outros veículos. Na categoria digital entram a inteligência artificial, a internet das coisas (OiT), *blockchain* e plataformas de economia sob demanda, como Airbnb e Uber.

Klaus Schwab, autor da obra “*A quarta revolução industrial*” (2018) enfatiza que da construção civil, até a advocacia, considerado um trabalho de cunho intelectual, a quarta revolução industrial impactará a todos. O autor adverte que não somente a indústria será impactada, mas também os governos. Nesse contexto, a Inteligência Artificial está em processo exponencial de crescimento, graças às estruturas globais organizadas em rede. As variadas conexões desta (ou destas) rede(s) estão fazendo emergir novas realidades, ou, talvez, um novo mundo. Alguns destes “nós” estão fortalecidos pela denominada “convergência tecnológica” de quatro setores, que são: nanotecnologia; biotecnologia; ciências cognitivas e tecnologia da informação, buscando atender às novas demandas e anseios do ser humano. (ENGELMANN; WERNER, 2020).

Engelmann e Werner (2020), consideram que este conjunto de tecnologias integram a quarta revolução industrial, fundamentada na expressão trazida por Klaus Schwab, que diz:

A Quarta Revolução Industrial é algo fabricado por nós mesmos e está sob nosso controle, e como as novas formas de colaboração e governança, acompanhadas por uma narrativa positiva e compartilhada, podem dar forma à nova Revolução Industrial para o benefício de todos. Se aceitarmos a responsabilidade coletiva para a criação de um futuro em que a inovação e a tecnologia servem às pessoas, elevaremos a humanidade a novos níveis de consciência moral. (ENGELMANN; WERNER, 2020).

Segundo Schwab, a Quarta Revolução Industrial apresenta três características:

Velocidade: ao contrário das revoluções industriais anteriores, esta evolui em um ritmo exponencial e não linear”, a evolução das diversas frentes da revolução está indo muito além da linearidade, mas estão progredindo de modo exponencial. (ENGELMANN; WERNER, 2020).

Amplitude e profundidade: ela tem a revolução digital como base e combina várias tecnologias, levando a mudanças de paradigma sem precedentes da economia, dos negócios, da sociedade e dos indivíduos. Aqui se verifica a passagem do analógico para o digital, mudança que deverá ser percebida e incorporada ao Direito e por todos aqueles que trabalham com o âmbito jurídico. Existe uma convergência transdisciplinar entre diversas áreas, produzindo o desenvolvimento exponencial, pois a revolução não é gerada por uma área do conhecimento, mas pela conjugação de diversos setores. (ENGELMANN; WERNER, 2020).

Impacto sistêmico: ela envolve a transformação de sistemas inteiros entre países e dentro deles, em empresas, indústrias e em toda sociedade. O sistema social e econômico como um todo, potencializados pelos efeitos da globalização, estão operando em conjunto e sendo movido pelas grandes possibilidades geradas pela revolução. (ENGELMANN; WERNER, 2020).

A nova transformação que ocorre da emergência de novas tecnologias no contexto da quarta revolução industrial aponta que os próprios computadores já sejam capazes de tomar decisões apropriadas e confiáveis. A automação das decisões relacionadas a vendas, por exemplo, mostra-se disruptiva em relação à era não digital, uma vez que permite que empresas usem o aprendizado das máquinas e a Inteligência Artificial para aproveitar grandes bases de dados, cruzando-as com dados disponíveis na Internet para criar listas prospectivas. Nesse contexto, sites prometem oferecer insight estratégico e

orientação tática na busca de pessoas para o tipo de negócio definido, com uma mensagem adequada<sup>4</sup>.

Com a grande conectividade, há em nossa era uma superexposição de informação, pois nunca a sociedade teve que lidar com tantos dados, de inúmeras fontes. Os autores também fazem um alerta em frente a esta realidade:

“Diante de tanta informação, a tecnologia que se sobressai é a inteligência artificial, sendo esta responsável por filtrar e tomar as melhores decisões diante do excesso de informação. Com a IA, novas tecnologias se reinventam para o avanço da civilização. Surge uma robótica totalmente conectada à internet, onde é controlada por uma inteligência artificial”. (TEIXEIRA; CHELIGA, 2021, p. 51)

Outra influência da IA na Quarta Revolução Industrial é a relação direta com a automação dos processos produtivos ao revesti-los de uma nova qualidade: a capacidade de aprendizado dos programas para melhorar o seu próprio desempenho. O processo de automação no setor produtivo já é bem conhecido e se mostra bastante difundido, seja na produção industrial de massa, seja na oferta de serviços. Assim, várias fábricas já produzem produtos – carros, por exemplo – com o uso intensivo de mão de obra robótica. O mesmo ocorre na prestação de serviços bancários e diversos outros. A diferença central da automação da Quarta Revolução Industrial se relaciona com a tentativa de produzir programas de computador que possam aprender. Um exemplo citado é o programa AlphaGo<sup>6</sup>, que houve por vencer um competidor daquele milenar jogo de tabuleiro chinês. O treinamento do programa teria ocorrido por meio de partidas simuladas contra ele mesmo. Assim, por meio de repetição, foi possível que o modelo algorítmico pudesse evoluir.

Os programas de computação relacionadas com a Quarta Revolução Industrial seriam centralmente construídos para essa finalidade: aprender e melhorar os modelos aplicados às atividades restritas para as quais tenham sido criados. O ponto central é que esses programas seriam especializados e restritos. O software AlphaGo é muito eficiente para modelar e

---

<sup>4</sup> SYAM, Niladri; SHARMA, Arun. Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice. *Industrial Marketing Management*, n. 69, p. 135-146, 2018.

jogar Go, mas não tem a mesma capacidade para modelar e jogar xadrez. Esse conceito é crucial para entender a automação da Quarta Revolução Industrial. Se as máquinas do passado e do presente seriam especializadas, as “máquinas” do futuro serão superespecializadas (VERONESE; SILVEIRA; LEMOS, 2020).

Fernanda de Carvalho Lage, na sua obra “Manual de Inteligência Artificial no Direito Brasileiro (2021)” acredita que não há dúvidas de que o período atual mostra um *boom* da IA. Porém, a autora indaga que algumas questões ainda pairam:

As promessas do que a IA é capaz não estão superestimadas? O desapontamento, por exemplo, com limitações de veículos sem motorista, processamento de linguagem natural (PLN) e análises preditivas baseadas em IA, não podem levar a uma queda do investimento, interesse e financiamento dessa tecnologia? (LAGE, 2021, p. 36)

Lage (2021, p. 36) destaca que à medida que a comunidade resolve e descobre mais problemas desafiadores se abrem. Dessa forma, a IA transforma-se em uma jornada de pesquisa científica de longo prazo. A IA será uma das tecnologias de ativação mais poderosa já criadas - uma única invenção que pode desbloquear soluções para milhares de problemas. A próxima década verá novos esforços para generalizar as capacidades dos sistemas de IA para ajudar a alcançar esse potencial- criando métodos que já foram bem-sucedidos e pesquisando como criar IA de uso geral que pode realizar uma ampla gama de tarefas.

Porém, o relatório de pesquisa do Grupo Gartner mostra que os gestores veem uma enorme oportunidade nos benefícios da IA, mas ainda lutam para colocar essas vantagens na prática. O Grupo Gartner foi criado por Gideon Gartner, no final da década de 70<sup>5</sup>, por ser uma das principais empresas mundiais especializadas pesquisa e consultoria em tecnologia da informação, possui a missão de gerar informações, métricas e análises a respeito de tecnologia para que seus clientes tomem decisões estratégicas. Possui 9.100

---

<sup>5</sup> Disponível em <https://rockcontent.com/br/blog/quadrante-magico-gartner/>. Acesso em: 23 dez. 2021

clientes, 6.600 associados, incluindo mais de 1.500 analistas de pesquisa, consultores, em 85 países<sup>6</sup>.

O relatório Gartner indica que a IA remodelará a forma como o trabalho é feito, pois a tecnologia substituirá algumas tarefas normalmente executadas pelo homem e mudará a forma como as decisões do dia a dia são tomadas. Atualmente, os casos de uso se enquadram principalmente em três categorias: automatização e otimização, geração de *insights* e criação de engajamento humano (por exemplo, chatbots e assistentes virtuais<sup>7</sup>).

No entanto, o relatório indica que falsas perspectivas sobre a IA trazem dificuldades para algumas organizações definirem as expectativas corretas em relação aos resultados do negócio. Essas falsas expectativas dão origem a projetos sem chance de sucesso. Quando isso acontece, conclui o relatório, os gestores, com expectativas irrealistas costumam culpar a tecnologia e a ciência por sua incapacidade de criar as transformações que esperavam. O relatório cita algumas falsas expectativas:

- IA e aprendizado de máquina (ML) são iguais e intercambiáveis;
- As máquinas inteligentes aprendem por conta própria;
- IA pode ser 100% objetiva;
- IA apenas substituirá empregos de atividades repetitivas;

A fim de mitigar futuras frustrações, o relatório sugere a criação de estratégias corporativa para IA, com vistas a identificar casos de uso e métricas de sucesso desde o início. As formas comuns de medir os benefícios incluem: redução de risco, velocidade do processo, maior satisfação do usuário e redução das necessidades ou custos de mão de obra. Por fim, concluem que muitos casos de negócios dependem de uma combinação de benefícios tangíveis e intangíveis.

O relatório Gartner 2021 também apontou a importância do tratamento dos dados no uso de IA. O relatório aponta que para obter o máximo

---

<sup>6</sup> Disponível em: <<https://blog.smartconsulting.com.br/quadrante-magico-gartner/>>. Acesso em 27 dez. 21

<sup>7</sup> Disponível em: <https://www.gartner.com/en/topics/artificial-intelligence>. Acesso em 23 dez. 21



impacto da Inteligência Artificial e do Aprendizado da Máquina, as organizações estão sendo obrigadas a implementarem soluções de IA mais “inteligentes”, com menor consumo de dados. A Gartner sugere que ao apoiar-se nesta tendência, as organizações aproveitarão os algoritmos de aprendizagem e os sistemas interpretáveis para reduzir o tempo de obtenção de valor e aumentar sua relevância comercial. Acredita-se que abordagens de dados pequenos e amplos fornecem análises robustas enquanto reduzem a dependência das organizações de grandes conjuntos de dados. Usando dados amplos, as empresas obtêm uma consciência situacional mais rica e completa ou uma visão de 360 graus, permitindo tomadas de decisão mais assertivas<sup>8</sup>.

A principal preocupação das instituições como a Gartner é a chamada *hype* da inteligência Artificial. *Hype* é algo exagerado e com uma conotação negativa. Está em *hype* qualquer assunto que esteja dando o que falar, que esteja na moda, mas que ao mesmo tempo não tenha fundamento comprovado. No contexto da tecnologia, algo que esteja em *hype* traz consigo um grande receio de que o estado atual da tecnologia não seja suficiente para solucionar os problemas que se propõe a enfrentar

Assim, o *hype* é consequência da nossa própria falta de domínio técnico, da nossa consequente superficialidade nesse campo. São ingredientes adicionais o interesse de pessoas de alimentarem o *hype*, por exemplo, um conferencista que reafirme seu suposto saber ou empresas que vendam o *hype*, pois elas funcionam na lógica de uma comunicação imediata e facilitada.

Os elementos finais são as palavras inteligência e artificial, que transmitem um sentido bastante equívoco daquilo que realmente são quando utilizadas em conjunto. Seria melhor que essa tecnologia não tivesse seu conteúdo induzido por vocábulos que achamos que compreendemos, pois fazem parte da nossa linguagem em outros contextos<sup>9</sup>.

---

<sup>8</sup> Disponível em: <<https://inforchannel.com.br/2021/03/29/gartner-identifica-as-10-principais-tendencias-em-data-analytics-para-2021/>>. Acesso em: 07 fev. 2022.

<sup>9</sup> Disponível em: <https://henarcos.com.br/depois-hype-inteligencia-artificial/>. Acesso em 21 jan. 2022.

## 1.2. Algoritmos

Algoritmo é um conjunto de etapas que, uma vez executadas, possibilitam a realização de uma tarefa. Algoritmos compostos por comandos em linguagens de programação são executados em computadores em todos os lugares – em laptop, em servidores, smartphone, em sistemas embutidos (como em carros, forno de micro-ondas ou em sistemas de ar-condicionado). Portanto, um algoritmo de computador é:

Um conjunto de etapas para executar tarefa descrita com precisão suficiente para que um computador possa executá-lo. Dada uma entrada para um problema, o algoritmo deve sempre produzir solução correta para este e usar recursos computacionais eficientemente ao fazê-lo. (CORMEN, 2017).

Em outras palavras, na computação, os algoritmos são as sequências de operações realizadas por uma máquina. Na essência, os algoritmos são apenas uma forma de representar matematicamente um processo estruturado para a realização de uma tarefa. É semelhante às regras e aos fluxos de trabalho, um passo a passo que ocorre nos processos de tomada de decisão em uma empresa, por exemplo<sup>10</sup>.

Lage (2021, p. 37) também define algoritmo como processo ou conjunto de regras a serem seguidas em cálculos ou outras operações de solução de problemas, especialmente por computadores. O objetivo de um algoritmo é resolver um problema específico, geralmente, definido por alguém como uma sequência de instruções.

Os algoritmos funcionam logicamente enquanto seguem as etapas para obter uma saída para uma determinada entrada. Um algoritmo é qualquer procedimento computacional bem caracterizado que toma algum valor, ou conjunto de valores, como entrada e produz algum valor, ou conjunto de valores, como resultado. Ele é uma sequência de etapas computacionais que transformam os dados de entrada em dados de saída<sup>11</sup>.

Na prática, os algoritmos proporcionam grande impacto na rotina das pessoas. Algoritmos determinam os *feeds* de notícias, as buscas na internet,

---

<sup>10</sup> ELIAS, Paulo Sá. Algoritmos, Inteligência Artificial e o Direito. Disponível em: <https://conjur.com.br/dl/algoritmos-inteligencia-artificial.pdf>. Acesso em: 17 de nov. 2020.

<sup>11</sup> Fonte: (CORMEN, Thomas H. et al. Introduction to algorithms. Cambridge: MIT Press, 2009.

as sugestões e influências de marketing digital, dentre outras coisas. Como boa parte da rotina atual das pessoas está baseada no uso de equipamentos eletrônicos, igual e proporcionalmente está o contato com algoritmos (HARTMANN, 2019).

Para que um algoritmo possa representar um processo, ele deve ser (LAGE, 2021):

- I. finito: o algoritmo deve eventualmente resolver problemas;
- II. bem definido: a série de etapas deve ser precisa e compreensíveis. Especialmente porque os computadores estão envolvidos em uso do algoritmo, o computador deve ser capaz de entender as etapas para crie um algoritmo utilizável.
- III. efetivo: um algoritmo deve resolver todos os casos do problema para os quais alguém o definiu. Um algoritmo deve sempre resolver o problema que foi criado para resolver.

Na construção, ele deve possuir as seguintes características descritas no quadro abaixo:

Tabela 1- Características do Algoritmo.

<b>Características do algoritmo</b>	<b>Descrição</b>
Entrada	Um algoritmo deve ter 0 entradas bem definidas.
Saída	Deve ter 1 ou saídas bem definidas e deve corresponder à saída desejada.
Viabilidade	Deve ser finalizado após o número finito de etapas.
Independência	Um algoritmo deve ter instruções passo a passo, independentes de qualquer código de programação.

Clareza	Deve ser inequívoco e claro. Cada uma de suas etapas e entradas/saídas devem ser claras e possuir apenas um significado.
---------	--

Fonte: LAGE (2021)

- I. Os algoritmos podem ser classificados em três tipos, conforme suas estruturas (CORMEN et al., 2009): sequência: tipo de algoritmo caracterizado por uma série de etapas, e cada etapa será executada uma após a outra.
- II. conexão: representado por problemas “se”. Se uma condição for verdadeira, a saída será A, se a condição for falsa, a saída será B. Esse tipo de algoritmo também é chamado de “tipo de seleção”.
- III. repetição: o processo pode ser executado repetidamente sob uma determinada condição. É representado pelos problemas “sim” e “não”. É necessário, contudo, que o processo termine após um certo número de loops, ou seja, repetições.

A inteligência artificial depende de algoritmos para suportar o seu funcionamento complexo, pois não se trata apenas de ciências da computação, mas também de matemática, lógica ou física. Portanto, diferentes técnicas são usadas para atingir a solução desejada. Lage (2021, p. 41) fundamentada nos estudos de Tim Roughgarden<sup>12</sup> (2017) classificou os algoritmos de acordo com sua função:

Tabela 2- Classificação dos Algoritmos

Classificação dos Algoritmos	Função
------------------------------	--------

<sup>12</sup> ROUGHGARDEN, Tim. Algorithms Illuminated: Part 1. São Francisco: Soundli Keyyourself Publishing, 2017. Disponível em: <http://algorithmsilluminated.org/>. Acesso em 20 dez. 2021.

Regressão e Classificação Linear	Simple e muito uteis quando a quantidade de atributos é muito grande.
Regressão Logística	Recurso que permite estimar a probabilidade associada à ocorrência de determinado evento em face de um conjunto de variáveis explanatórias.
Árvore de decisão	Geralmente similar ao processo de decisão das pessoas e é fácil de interpretar. Entretanto, é mais utilizado em combinação.
K-médias	É um algoritmo de agrupamento de dados não hierárquico que utiliza uma técnica iterativa para particionar um conjunto de dados.
PCA	Provavelmente, o método multivariado mais usado e antigo de redução de dados inventado por Pearson (1901) e Hotelling (1933).
Redes Neurais	Representa a nova leva de algoritmos de aprendizado de máquina e podem ser aplicados em muitas tarefas. Todavia, apresentam uma dificuldade por necessitarem de máquinas com grande capacidade de processamento computacional.

Fonte: LAGE (2021)

O professor Fabiano Hartmann, na sua obra “Inteligência Artificial e Direito” (2019), aponta uma importante distinção entre algoritmos e programa computacional. O programa não precisa satisfazer o *finiteness*<sup>13</sup>. Ainda, um algoritmo pode ser descrito com várias formas, desde a linguagem natural até por representações gráficas em um fluxograma, objetivando sempre

<sup>13</sup> No rastreamento de instruções de um algoritmo, para todos os casos, o algoritmo terminará após um número finito de etapas. Disponível em: <https://www.cmpe.boun.edu.tr/~akin/cmpe223/chap1.htm>. Acesso em 30 de julho de 2021.

o caminho mais objetivo e simples, para preservar suas características essenciais.

Hartmann (2019, p. 73) traz também a reflexão sobre a transparência algorítmica, onde ele ensina que na área da Ciência da Computação que busca a descoberta automatizada de correlações e padrões para estabelecer tomadas de decisão, pode-se chegar à possibilidade de se exercer uma espécie de poder (em campos proibidos ou, ao menos não desejáveis) de forma dificilmente identificável. A transparência, por sua vez, é um conceito poderoso. Contudo, a pura exposição algorítmica, a pura e simples e simples disponibilidade de códigos-fonte ou auditorias irrefletidas podem gerar uma ilusão de clareza.

Dessa forma, é importante a reflexão sobre como regular a aplicação algorítmica em softwares, tanto públicos como privados, estimular mecanismos de detecção de erros e aplicações inadequadas, numa especialização de sistemas de governança (HARTMANN, 2019).

Cathy O' Neil (2016, p. 30, tradução nossa), respeitada matemática norte-americana, em sua obra *Weapons of Math Destruction* (Arma de Destruição Matemática), alerta para os perigos e consequências da obscuridade dos códigos algorítmicos de IA, e as possíveis consequências desse procedimento.

Muitas empresas se esforçam para esconder os resultados de seus modelos ou mesmo sua existência. Uma justificativa comum é que o algoritmo constitui um “molho secreto” crucial para seus negócios. É propriedade intelectual e deve ser defendida, se necessário, com legiões de advogados e lobistas. No caso de gigantes da web como o Google, Amazon e Facebook, somente esses algoritmos precisamente adaptados valem centenas de bilhões de dólares. WMDs (Armas de destruição matemática) são, por concepção, caixas pretas inescrutáveis. Isso torna ainda mais difícil responder definitivamente a segunda pergunta: o modelo funciona contra o assunto interesse? Em suma, é injusto? Danifica ou destrói vidas? (CATHY O' NEIL, 2016, p. 30, tradução nossa).

## 1.3. Modelos de inteligência artificial

### 1.3.1. Machine learning

O *Machine Learning* é a área da IA responsável por grande parte das conquistas do campo nos últimos anos. O desenvolvimento da estrutura de *hardware*- dos processadores mais especificamente- e a disponibilidade de grande volume de dados possibilitaram que se colocassem em prática ideias que remontam ao início da computação (HARTMANN, 2019).

Lage (2021, p. 71) define *machine learning* como um subconjunto da IA e um superconjunto do aprendizado profundo de máquina (*deep learning*). “Em geral, o objetivo do aprendizado de máquina é entender a estrutura dos dados e integrá-los a modelos que possam ser entendidos e usado por todos.”

O *machine learning* também pode ser definido como o processo de desenvolver um problema prático por (LAGE, 2021):

1. Reunir um conjunto de dados;
2. Construir algoritmicamente um modelo estatístico com base nesse conjunto de dados;

Por meio do treinamento do modelo, o problema é resolvido. Habitualmente, um modelo de aprendizado de máquina (AM) deve ser submetido a uma fase de treinamento com um conjunto de dados suficientes para que ele “aprenda” para depois poder ser utilizado. Outro desafio para a IA, especificamente para a aplicação de modelos de AM, são os cenários em que há uma ausência total ou insuficiente de dados para serem utilizados na fase de treinamento dos modelos. Por exemplo, quando se quer implantar um sistema de AM em um sistema de informação recém implantado. Essas situações provocaram um desafio na área de IA, qual seja, oferecer modelos de AM que sejam capazes de aprender paulatinamente com a apresentação de novos dados – assim como os seres humanos o fazem (SILVA, 2021).

Na contemporaneidade, um elemento definidor dos sistemas de Inteligência Artificial é a multiplicidade de formas, a partir das quais se dá a habilidade de aprendizado pelas máquinas com base em informações passadas. Podem-se identificar três tipos de processos de aprendizado primordiais: o supervisionado, o não supervisionado e o aprendizado reforçado. O primeiro mapeia um conjunto de inputs para um dado conjunto de resultados, incluindo métodos como regressão linear, árvores de classificação e redes neurais. No segundo, os inputs são rotulados, mas os resultados não, o que significa que o algoritmo precisa inferir a estrutura subjacente dos próprios dados, como na análise de clusters, que visa a agrupar elementos em categorias similares, mas nas quais nem a estrutura dos clusters nem seu número são conhecidos antecipadamente. Os usuários precisam colocar uma confiança maior no sistema de Inteligência Artificial. Isso ocorre, por exemplo, no reconhecimento do comando de voz dado pelo usuário. No aprendizado reforçado, o sistema recebe um resultado variável para ser maximizado e uma série de decisões que podem ser tomadas para impactá-lo<sup>14</sup>.

### *1.3.1.1. Aprendizado supervisionado*

Nesse tipo de aprendizado, a máquina é ensinada pelo exemplo. É fornecido ao algoritmo de aprendizado de máquina um conjunto de dados conhecido que inclui as entradas e saídas desejadas, e o algoritmo deve encontrar um método para determinar como chegar a essas entradas e saídas. O algoritmo identifica padrões nos dados, aprende com a observação e faz previsões, e é corrigido pelo operador, e esse processo continua até que o algoritmo atinja um alto nível de precisão e desempenho (LAGE, 2021).

Esse tipo de aprendizado costuma ser usado na realização de tarefas de classificação (predição de uma classe) e de regressão (previsão de um valor numérico contínuo). Há a necessidade de um conjunto de dados rotulados- mas cujo rótulo não será submetido para a máquina- que será o

---

<sup>14</sup> KAPLAN, Andreas; HAENLEIN, Michael. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, v. 62, n. 1, p. 15-25, 2019.



*dataset* de teste e, finalmente, uma etapa de validação do desempenho da máquina em comparação aos rótulos esperados (PEIXOTO; SILVA, 2019).

Assim, a máquina é treinada usando dados bem “rotulados”, marcados. É um modelo mais simples que os outros tipos. A principal vantagem do aprendizado supervisionado é que ele permite a seleção e coleta dos dados, bem como a produção de uma saída de dados da experiência anterior. A ideia principal deste modelo é: os dados de entrada fornecem exemplo de situações e, para cada exemplo, é informado à máquina um resultado. Em seguida a máquina usará os dados de treinamento (*training data*) para construir o modelo que pode prever o resultado dos novos dados com base nos exemplos anteriores (LAGE, 2021).

A tarefa de classificação é a forma de *machine learning* usada de forma mais ampla para resolver problemas interessantes e difíceis da vida real. Entre algumas aplicações, destaca-se: classificação de documentos, aplicação de filtros de spam em e-mails, classificação de imagens, reconhecimento de palavras escritas a mão, detecção e reconhecimento de rostos em fotos. A precisão da classificação ou da regressão depende da efetividade do algoritmo escolhido, de como ele é aplicado, da qualidade e quantidade dos dados para treinamento.

### 1.3.1.2. *Aprendizagem Não Supervisionada*

O *machine learning* não supervisionado consegue dispensar a calibragem inicial dos seres humanos. Isso é alcançado por meio do desenvolvimento de novas tecnologias, como as redes neurais ou o *deep learning*, que são capazes de criar padrões de correlações próprios, alheios ao raciocínio humano. Isso é alcançado por meio da criação de uma rede de múltiplas unidades não lineares de processamento de dados que se retroalimentam de modo a imitar um cérebro humano. Esses sistemas de IA são capazes de analisar um ambiente dinâmico e dele extrair correlações e padrões. (GUTIERREZ, 2020).

O aprendizado não supervisionado costuma ser chamado de descoberta do conhecimento (*knowledge Discovery*). É um problema menos definido, pois não se sabe que tipo de padrões se está procurando e não há uma métrica clara de erro para controle- ao contrário do aprendizado supervisionado, em que se sabe o tipo de padrão almejado e a métrica adotada (PEIXOTO; SILVA, 2019).

Assim, o algoritmo estuda os dados para identificar padrões, não há programador humano para fornecer instruções. Ao invés disso, a máquina determina correlações e relacionamentos analisando os dados disponíveis.

Esse modelo é aquele em que só são conhecidos os dados de entrada, mas ainda não se sabe quais são as variáveis de saída correspondentes. Diferentemente do aprendizado supervisionado, não há um treinamento de dados feito por humanos: os algoritmos são deixados por conta própria para descobrirem e apresentarem a estrutura mais interessante nos dados (LAGES, 2021).

Lage (2021, p. 60) conclui que nesse tipo de aprendizado, é feita uma tentativa de determinar a estrutura dos dados, ou seja, tenta-se agrupar grandes quantidades de dados com base na característica que compartilham. A aprendizagem não supervisionada recebe esse nome porque nada é conhecido. É uma maneira de encontrar hierarquia e ordem em um conjunto de dados não estruturado. Esse modelo também possui a característica de realizar suas operações em tempo real, para que todos dados de entrada sejam analisados e rotulados na presença dos usuários, além de que os dados são rotulados de um computador são mais fáceis de serem obtidos do que os rotulados, que necessitam de intervenção manual.

### *1.3.1.3. Aprendizado por esforço*

O aprendizado por esforço é provavelmente o sistema de aprendizado mais complexo, uma vez que exige que a máquina seja equipada com sistemas e ferramentas capazes de melhorar o próprio aprendizado e integrar as características do ambiente no seu retorno (LAGE, 2021).

Alguns algoritmos de *machine learning* não têm sua experiência restrita a um *dataset*. É o caso da aprendizagem por esforço, que interagem com um ambiente no qual há um elo entre sistemas e suas experiências. Este tipo de aprendizagem explora uma abordagem computacional para o aprendizado pela interação. Ao invés de teorizar sobre como as pessoas aprendem, adota-se a perspectiva de um pesquisador de IA. Explorando-se situações idealizadas de aprendizagem para se avaliar a efetividade de uma série de métodos, explorando concepções que resolvam problemas com interesse científico ou econômico, avaliando essas concepções por análises matemáticas e experimentos computacionais (PEIXOTO; SILVA, 2019).

O professor Hartmann (2019, p. 95), fundamentado nos estudos de Sutton e Barto (2017), ensina que o aprendizado por reforço se difere do supervisionado, pois o último ocorre a partir de uma amostra de dados com exemplos rotulados por um supervisor externo com conhecimento na área. Cada exemplo descreve uma situação junto da especificação (o rótulo) da ação correta que o sistema deve tomar naquela situação. O objetivo desse tipo de aprendizado é que o sistema extrapole suas reações para que consiga agir corretamente em situações que não estavam sujeitas ao conjunto de treinamento. Essa forma de treinamento é de fato muito importante, mas que sozinha não é adequada para aprendizagem a partir da interação. Hartmann (2019) completa que não é prático obter exemplos de comportamentos desejados que sejam ao menos corretos e representativos das situações que o agente tem de agir em problemas interativo. O território desconhecido é onde se espera que o aprendizado seja mais proveitoso, pois o agente é obrigado a aprender por sua própria experiência.

O aprendizado por esforço também se difere do aprendizado não supervisionado, que tipicamente se resume em encontrar estruturas escondidas em conjunto de dados não rotulados. Hartmann (2019, p. 96) afirma que os termos supervisionado e não supervisionado podem parecer exaustivos para a classificação dos paradigmas de *machine learning*, mas não são. Ainda que a ideia de se pensar que a aprendizagem por esforço seja um tipo de aprendizagem não supervisionado por não contar com exemplos de comportamento esperado, o aprendizado por esforço busca maximizar um sinal

de bom desempenho, e não encontrar uma estrutura oculta- pode até se mostrar útil no aprendizado por esforço, mas por si só não dá conta do problema de maximizar um sinal de recompensa. Assim, o autor considera, por tanto, a aprendizagem por esforço, um terceiro paradigma do aprendizado de máquina.

Lage (2021, p. 79) conclui que o aprendizado por esforço é provavelmente o sistema de aprendizado mais complexo, uma vez que exige que a máquina seja equipada com sistemas e ferramentas capazes de melhorar o próprio aprendizado e, acima de tudo, integrar as características do ambiente no seu retorno. A máquina será alimentada com uma série de elementos de suporte, como por exemplo sensores, câmeras, GPS, permitindo que verifique o que está acontecendo no ambiente a sua volta e faça escolha para uma melhor adaptação ao ambiente ao seu redor. Esse tipo de aprendizado é encontrado, por exemplo, em carros não tripulado, ao qual por meio de um complexo sistema de sensores de suporte são capazes de percorrer estradas urbanas e não urbana reconhecer obstáculo, seguir placas e muito mais.

### 1.3.2. REDES NEURAIS E DEEP LEARNING

Aprendizado profundo é o subcampo da inteligência artificial que se concentra na criação de grandes modelos de rede neural que são capazes de tomar decisões baseadas em dados precisos. O aprendizado profundo é especialmente adequado para contextos em que os dados são complexos e onde existem grandes conjuntos de dados disponíveis. Hoje, a maioria das empresas online e de tecnologias de consumo de ponta usam aprendizado profundo<sup>15</sup>.

O aprendizado profundo é considerado a vanguarda da pesquisa em aprendizado de máquina. Consiste em uma abordagem algorítmica baseada no conceito de redes neurais. Parte-se da inspiração da compreensão da biologia do cérebro humano- as interconexões entre os neurônios. Mas, diferentemente de um cérebro humano no qual um neurônio pode se conectar a

---

<sup>15</sup> Kelleher, John D, Deep Learning, The MIT Press, 2019

qualquer outro neurônio a uma certa distância física, essas redes neurais artificiais possuem camadas, conexões e direções distintas de propagação de dados (LAGE, 2021).

Apesar dos estudos recentes e promissores sobre o aprendizado profundo de máquina, Peixoto e Silva (2019, p. 97) esclarecem que as redes neurais artificiais fazem parte do mundo da computação há muito tempo. Seu uso foi chamado de *cybernetics* (1940s-1960s), *connectionism* (1980s-1990s), até que foi referenciado como *deep learning* no início dos anos 2000, quando as redes neurais começaram a se aprofundar.

O Professor Hartmann (2019, p. 98) define rede neural como um tipo de sistema computacional inspirado pelas propriedades básicas de neurônios biológicos. Uma rede neural é composta por muitas unidades individuais, conectadas por ligações direcionadas, cada qual recebe um *input* de umas e envia *outputs* a outras, propagando a ativação da rede.

As explicações matemáticas do funcionamento das redes neurais são bastantes complexas. De forma simplista, porém, é possível colocar as redes neurais como a forma pela qual ocorre o aprendizado de máquina profundo: *deep learning*. Essas estruturas são projetadas pelos programadores para melhor atender ao problema proposto e podem variar os seus pesos, número de camadas e tamanho (PEIXOTO; SILVA, 2019).

Assim, Lage (2021, p. 82) apresenta diferenças cruciais entre aprendizado de máquina (*machine learning*) e aprendizado profundo de máquina (*deep learning*):

- a) aprendizado de máquina (*machine learning*) usa algoritmos para analisar todos dados, aprender com esses dados e tomar decisões informadas com base no que aprendeu;
- b) aprendizado profundo de máquina (*deep learning*) estrutura algoritmos em camadas para criar uma rede neural artificial, uma tentativa de reproduzir o trabalho do cérebro humano em um computador usando camadas de neurônios, que pode aprender e tomar decisões inteligentes por conta própria;

- c) o *deep learning* é um subconjunto do *machine learning*. Entretanto, ambos se enquadram na ampla categoria de IA, o primeiro é o que alimenta a inteligência artificial mais semelhante à humana.

Modelos de IA, diferentemente dos seres humanos, ainda não conseguiram generalizar o conhecimento de uma forma suficientemente eficaz para a solução de uma situação nova. Pesquisas em IA que têm buscado algoritmos para viabilizar esta possibilidade (Aprendizado Contínuo) são identificadas pelos termos *Lifelong Learning Algorithm* (LLA) ou *Lifelong Machine Learning* (LML) (SILVA, 2021).

Outros trabalhos buscam a operacionalização de Aprendizado Contínuo e têm adotado a estratégia de Algoritmos de Aprendizado Incremental, Retreinamento Manual ou Retreinamento Automático. A primeira abordagem é dependente de modelos de aprendizado de máquina (AM) que apresentam características de aprendizado genuinamente incremental, ou seja, sua assertividade melhora, conforme novos dados lhes são apresentados em uma dinâmica concomitante de uso e aprendizado. Essa estratégia é a que mais se assemelha ao comportamento humano. No entanto, geralmente, modelos desse grupo apresentam performance preditiva aquém dos modelos tradicionais. As outras duas estratégias propõem que se apliquem reiterados treinamentos em modelos tradicionais de forma a perpetuar suas performances preditivas (SILVA, 2021).

No retreinamento manual, há de se contar com um ser humano que faça novos treinamentos na medida em que o modelo diminui sua performance preditiva. Essa situação é muito semelhante ao treinamento tradicional e, por via de regra, restaura a boa performance preditiva do modelo. Por outro lado, pode ser onerosa, caso os novos treinamentos tenham de ser em intervalos curtos – consumindo muito tempo de trabalho de um especialista em IA.

No caso de Retreinamento Automático, há de se definir um sistema paralelo capaz de checar a performance preditiva e proceder com novos treinamentos sempre que a performance preditiva ficar abaixo de um limiar. Nesse caso se garante a boa performance preditiva do modelo continuamente e

de forma automática. O desafio está em se definir o sistema paralelo de checagem de performance preditiva (SILVA, 2021).

### 1.3.3. PROCESSAMENTO DE LINGUAGEM NATURAL (PLN)

Segundo Vieira e Lopes (2010), Processamento de Linguagem Natural (PLN) é uma área de Ciência da Computação que estuda o desenvolvimento de programas de computador que analisam, reconhecem e/ou geram textos em linguagens humanas, ou linguagens naturais.

Glavas et al (2012) ensinam que tarefas de Processamento de Linguagem Natural, como classificação e resumo de texto, recuperação de informação e desambiguação de sentido de palavras, dependem de uma medida de semelhança semântica de documentos textuais.

Manning e Schütze (2000) analisam que uma abordagem estatística de PLN procura resolver esses problemas automaticamente aprendendo preferências lexicais e estruturais dos corpora. Em vez de analisar apenas usando categorias sintáticas, como parte de rótulos de fala, reconhecemos que há muita informação nas relações entre as palavras, ou seja, quais palavras tendem a se agrupar entre si. Este conhecimento pode ser explorado como uma janela para relacionamentos semânticos mais profundos. Em particular, o uso de modelos estatísticos oferece uma boa solução para o problema da ambiguidade: os modelos estatísticos são robustos, generalizam-se bem e comportam-se graciosamente na presença de erros e novos dados.

Assim, os métodos de PLN estatístico abriram o caminho para fornecer uma desambiguação bem-sucedida em sistemas de larga escala usando textos que ocorrem naturalmente. Além disso, os parâmetros dos modelos estatísticos de PLN podem frequentemente ser estimados automaticamente a partir de corpora de texto, e essa possibilidade de aprendizado automático não apenas reduz o esforço humano na produção de sistemas de PLN, mas também levanta questões científicas interessantes sobre a aquisição da linguagem humana.

### 1.3.4. CORPUS

Para Sinclair (2005), um corpus é uma coleção de peças de linguagem em formato eletrônico, selecionados de acordo com critérios externos para representar, na medida do possível, uma linguagem ou variedade linguística como fonte de dados para a pesquisa linguística.

Os vetores de palavras representam um salto significativo no avanço de nossa capacidade de analisar relacionamentos entre palavras, sentenças e documentos. Ao fazê-lo, eles avançam a tecnologia fornecendo às máquinas muito mais informações sobre as palavras do que anteriormente era possível usando representações tradicionais. São vetores de palavras que possibilitam tecnologias como reconhecimento de fala e tradução automática (MARTINS, 2018).

Abordagens tradicionais da PLN, como os modelos de codificação e de saco de palavras quentes, embora úteis em algumas tarefas de aprendizagens de máquina, não capturam informações sobre o significado ou contexto de uma palavra. Isso significa que relacionamentos potenciais, como a proximidade contextual, não são capturados em coleções de palavras. Por exemplo, uma codificação simples não pode capturar relações simples, como determinar se as palavras “cachorro” e “gato” se referem a animais que são frequentemente discutidos no contexto de animais domésticos (MARTINS, 2018). Essas codificações geralmente fornecem linhas de base suficientes para tarefas simples de PLN (por exemplo, classificadores de *spam* de *e-mail*), mas não possuem a sofisticação para tarefas mais complexas, como tradução e reconhecimento de fala. Em essência, as abordagens tradicionais da PLN, como as codificações únicas, não capturam relações sintáticas (estrutura) e semânticas (significado) entre coleções de palavras e, portanto, representam a linguagem de uma maneira muito ingênua. (AHIRE, 2018)

O modelo de IA do Athos utiliza termos que acompanham um termo alvo para inferir seu significado, assumindo que sinônimos são, geralmente, usados no mesmo contexto. A teoria corresponde ao famoso adágio: ‘diz-me com quem tu andas e lhe direi quem tu és’. Assim, se dois termos habitualmente aparecem acompanhados de um grupo de outros termos, é



provável que esses dois termos possuam uma relação semântica, desde que observadas em um mesmo contexto. Observemos os termos “areia” e “praia”, sabidamente não são sinônimos, mas é possível observar que as sentenças “descansar na praia” e “descansar na areia” possuem uma grande proximidade semântica (MARTINS, 2018).

Goldberg (2017) explica que embora a semelhança de palavras seja difícil de definir e geralmente seja muito dependente de tarefas, as abordagens atuais derivam da hipótese distributiva [Harris, 19547], afirmando que as palavras são semelhantes se aparecerem em contextos semelhantes.

### 1.3.5. Paragraph Vector

Para Martins (2018), o conceito por trás do *Paragraph Vector* é basicamente uma expansão do modelo Word Vector para que opere com parágrafos ou documentos inteiros.

Numa definição mais formal, Le & Mikolov<sup>16</sup> (2014) explicam a proposta como uma estrutura não supervisionada que aprende representações contínuas de vetores distribuídos para partes de textos. Os textos podem ser de tamanho variável, variando de sentenças a documentos. O nome *Paragraph Vector* enfatiza o fato de que o método pode ser aplicado a textos de tamanho variável, desde uma frase ou sentença até um documento grande. O modelo *Paragraph Vector* tem sido utilizado nas mais diversas áreas do processamento de linguagem natural com muito sucesso (MARTINS, 2018).

---

<sup>16</sup> LE, Quoc V. MIKOLOV, Tomas: “Distributed Representations of Sentences and Documents”, Mountain View, CA: Google Inc. 2014.

## 2. Inteligência artificial no Poder Judiciário

### 2.1. Inteligência Artificial e decisões judiciais

O tema inteligência artificial e direito, especificamente a IA e as decisões judiciais, é atual e está na pauta de estudo e desenvolvimento em vários países, assumindo centralidade nos principais espaços públicos de debate. Atualmente, a IA tornou-se um centro de atenção de profissionais do direito, isso se dá por inúmeros fatores: avanços tecnológicos das ferramentas da IA incluindo, entre elas, as já citadas aprendizado de máquina (*machine learning*), aprendizado profundo (*deep learning*) e processamento de linguagem natural; a mudança de atitude dos profissionais da área jurídica em relação à tecnologia; a grande disponibilidade de dados na rede e o advento do processo judicial eletrônico; o sucesso recente de aplicativos de IA, *Law Tech*, no domínio público privado (LAGE, 2020).

Nesse sentido, Engelmann e Werner (2020) destacam a importância do Direito em resolver lides e viabilizar a tomada de decisões no sentido de produzir soluções, sustentada em regramentos previamente conhecidos. Entretanto, os autores, acreditam que esse cenário está passando por rápidas transformações, através do surgimento de sistemas, de algoritmos, viabilizadores da inteligência artificial desenvolvidos pelo próprio ser humano para tomar decisões, avançando das mais simples às mais complexas. Ao mesmo tempo, o Direito, como área de conhecimento, sempre esteve assentado nos pressupostos da certeza, segurança e previsibilidade.

Embasados nos estudos de Ulrich Beck<sup>17</sup>, os autores acreditam que tudo isto está em transformação, caracterizado pela chamada “metamorfose do mundo”: ou seja, “a metamorfose desestabiliza as certezas da sociedade moderna, os eventos e processos que provocam um choque fundamental; [...] a metamorfose significa que o que foi impensável ontem é real e possível hoje”. Beck (2018) conseguiu sintetizar as possibilidades já evidenciadas de que a

---

<sup>17</sup> BECK, Ulrich. A metamorfose do mundo: novos conceitos para uma nova realidade. Trad. Maria Luiza X. de A. Borges. São Paulo: Zahar, 2018. p. 10-11.

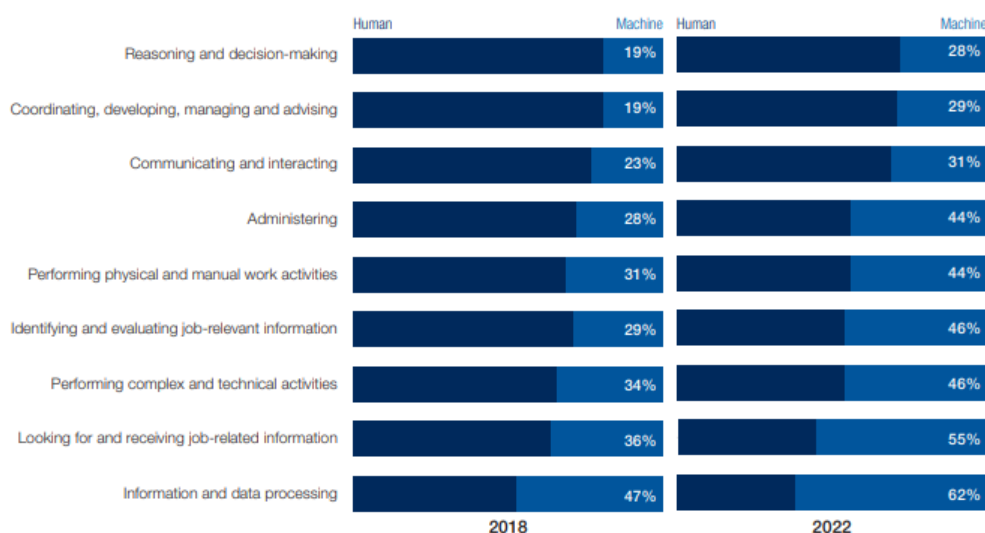
inteligência poderá trazer, permitindo a construção de redes, relações, isto é, um mundo que se apresenta muito diferente do que se estava acostumado a vislumbrar. O próprio Beck completa:

A metamorfose implica uma transformação muito mais radical, em que as velhas certezas da sociedade moderna estão desaparecendo e algo inteiramente novo emerge”. Este “novo” que está em fase de nascimento, pode ser comparando a um iceberg, onde uma parte está desvelada e uma grande porção ainda não é visível, que traz um misto de curiosidade e ansiedade, dado o desconhecido desta face ainda submersa. (BECK, 2018, p. 10-11).

Portanto, as diversas carreiras jurídicas terão o desafio de observar o cenário, a fim de se perceber quais são as novas necessidades do mercado e as faixas remanescentes, após o ingresso da inteligência artificial. Será preciso ocorrer uma reinvenção do papel do jurídico no encaminhamento da solução para os problemas sociais, além de se promover uma modernização das fontes do Direito e do modo como se atribui efeitos jurídicos aos fatos sociais (ENGELMANN; WERNER, 2020).

Nesse sentido, alguns estudos já apontam para a transformação da relação entre horas de trabalho humano/máquina, conforme o documento intitulado: *The Future of Jobs Report 2018*, elaborado e publicado pelo Fórum Econômico Mundial:

Gráfico 1- Transformação das carreiras.



Através da análise, realizada por Engelmann e Werner (2020), dos dados apresentados pelo Fórum Econômico Mundial, verifica-se um crescimento da participação da máquina, da automação e da inteligência artificial, quando comparamos a quantidade de horas utilizados pelos humanos em relação às máquinas. Os autores deram maior atenção a alguns itens específicos: nas horas dispendidas entre humanos e máquinas para “raciocínio e tomada de decisão” se observa que, em 2018, a máquina usa 19% do total, quantidade que está projetada, para 2020, a passar para 28% do total das horas gastas neste quesito. Apesar deste aumento, a participação humana, na projeção para 2020, ainda fica em 72%. De qualquer modo, representa um aumento projetado de 9% para um período de 2 anos.

Embora seja uma categoria bem vaga, ela poderá, no caso da advocacia, corresponder à pesquisa doutrinária, legislativa e jurisprudencial. Nesse caso, passa-se de 29% do total das horas medidas em 2018 para uma projeção de 46%, em 2020, da carga horária realizada pela máquina. Aqui se verifica o incremento mais significativo, ou seja, 17%, e talvez seja justamente, pelo que se pesquisou até o momento, o setor onde a inteligência artificial esteja ingressando com mais força. Aqui se observa que a participação humana nesta categoria de atividades ficará, no projetado para 2020, com 54% da carga horária gasta.

Na parte referente à “realização de atividades complexas e técnicas”, observa-se um crescimento menor, mas que não deverá ser desprezado, pois se passa de 34% das horas gastas com esta atividade e realizada pela máquina para 46% do total das horas gastas pela máquina. Se observa um crescimento de 12% em um período de dois anos, ou seja, de 2018 para a projeção de 2020.

O último quesito analisado se refere à “informação e processamento de dados”, onde se observa que, em 2018, do total das horas gastas com este item, 47% já é realizado pela máquina; e se faz uma projeção para 2020 de 62%. Vale ressaltar, um incremento de 15%.

Esse é o panorama que se tem até o momento. Resta saber como a área do Direito irá lidar com estas projeções do incremento da inteligência artificial. Trata-se de um alerta. Por ser uma projeção, há uma probabilidade de que os dados venham a se confirmar, ou não. Dessa forma, está-se em meio a uma encruzilhada e que exigirá uma tomada de decisão que se lança do presente para o futuro. (ENGELMANN; WERNER, 2020).

Billion e Guilhermin (2018) produziram duas abordagens que contribuem para a informatização do direito: a primeira vem da modelagem do conhecimento jurídico usando sistemas baseados em regras, ou em sistemas especialistas, dentro da estrutura de algoritmos determinísticos, a fim de relatar a articulação lógica de certas regras legais. “Um país regido pela *Civil Law*, como Brasil, cujas leis estão previstas em códigos, constitui a priori, terreno favorável para esse tipo de abordagem” (LAGE, 2021). A segunda, mas adaptáveis aos países da *Common Law* que adotam o sistema de precedentes judiciais, conta com ferramentas de processamento que permite explorar grandes quantidades de documentos, como decisões judiciais digitais, para identificar automaticamente as respostas geralmente fornecidas para aquela controvérsia.

Os dados apresentados pelo Fórum Econômico Mundial, juntamente com as abordagens apresentadas por Billion e Guilhermin (2018), nos remete a nova tendência incorporada ao direito: a Jurimetria, que pode ser genericamente definida como “a estatística aplicada do Direito”. A Jurimetria parte da premissa de que o Direito não se esgota no estudo teórico das leis. Nunes (2016) nos alerta sobre a necessidade de estudarmos também os processos de decisão, através de todas as normas, gerais e individuais formuladas:

A Jurimetria também assume que este estudo precisa ser concreto, ou seja, deve situar seu objeto no tempo e no espaço e investigar os principais fatores capazes de interferir nos seus resultados. E a Jurimetria acredita que o estudo dos processos de decisão deve abandonar pretensões deterministas e admitir no ambiente acadêmico aquilo que sempre se admitiu no ambiente profissional: que a complexidade da ordem jurídica não permite afirmações absolutas e que o Direito, como tudo que envolve a vontade humana, é variável e incerto. Portanto, compreender o Direito é, antes de tudo, descrever suas variabilidades e controlar (não extirpar) suas incertezas. (NUNES, 2019).

Nunes (2016) acredita que o movimento empírico no Direito está avançando rapidamente. Todas as grandes Escolas dos EUA querem contratar juristas e acadêmicos empíricos. A Conferência de Estudos Empíricos Jurídicos tem cada vez mais audiência. A Revista de Estudos Jurídicos Empíricos está prosperando, bem como seus competidores. Cada vez mais, o melhor da Academia Jurídica é publicado em revistas especializadas e muito deste trabalho é empírico. E como toda mudança de referência, o surgimento do empirismo jurídico (leia-se: quantitativo) tem mais de uma causa. A tecnologia certamente não é a mais importante. Algumas técnicas estatísticas que indiscutivelmente são convenientes às questões de pesquisa jurídica requerem computação poderosa: análise bayesiana, modelos de equação estrutural e *bootstrapping*<sup>18</sup> são exemplos.

A quase ubíqua disponibilidade de computadores possibilita a compilação de grandes bases de dados, que se originam diretamente do sistema judiciário. Alguns países têm deslocado a troca de documentos no julgamento para interação computadorizada. Então, o material bruto do trabalho empírico já estará disponível num formulário eletrônico. Ainda melhor se esse material bruto não estiver apenas em imagens de texto, mas já codificado e pronto para análise. O Brasil está na vanguarda desse desenvolvimento, pois possui base de dados de amplitude e profundidade únicas, sem precedentes. (NUNES, 2016).

Os indicadores de ordem qualitativa que deveriam ser objeto de um debate ético pelos gestores judiciais, de forma a se estruturar uma estatística judicial que efetivamente passasse a se aproximar da realidade qualitativa da instituição, em sua missão de fazer justiça. O incremento da possibilidade de simples outorga de soluções formais, a depender do emprego que dê à inteligência artificial, torna ainda mais premente a revisão das balizas hoje empregadas para a estatística judiciária. (RODRIGUES, 2021).

## 2.2. Inteligência artificial no Poder Judiciário

As estatísticas indicam que o Brasil possui um dos maiores estoques de processos judiciais do mundo, com aproximadamente 79 milhões

---

<sup>18</sup> *Bootstrap* significa criar sua startup usando somente recursos próprios, não recorrendo a investidores externos. Disponível em: [www.sebrae.com.br](http://www.sebrae.com.br). Acesso em: 09. ago. 2021

de processos, e um elevado índice de congestionamento, cerca de 70%, apesar de contarmos com cerca de 18 mil juízes<sup>19</sup>. Essa realidade, além de dificultar o acesso à Justiça e ampliar a duração dos processos, tem também elevado significativamente o custo e a manutenção do sistema de justiça, que segundo dados do Conselho Nacional de Justiça (CNJ), corresponde a cerca de 1,4% do Produto Interno Bruto (PIB). O atual estágio da sociedade da informação demanda a modernização do Estado, não apenas para fazer frente a esse número expressivo de demandas, como também para possibilitar a mineração dos dados estratégicos para o planejamento da gestão. Nesse contexto, o uso da tecnologia não é apenas uma realidade, mas uma necessidade. (CUEVA, 2021).

A ideia de auxílio ou de alocação isolada de soluções baseadas em IA no Judiciário não é novidade. No fim do século passado, Sartor e Branting (1998) sintetizaram importantes possibilidades de aplicação da IA para o processo de tomada de decisão no Judiciário.

Os autores já identificaram, à época, estudos que abordavam rotinas com procedimentos que envolviam aplicações diretas com regras objetivas a fatos incontroversos e pesquisas concentradas em áreas complexas e desafiadoras do processo legal, tipificadas pela tomada de decisões judiciais<sup>20</sup>.

A tomada de decisões judiciais é área de complexidade assustadora, em que o conhecimento jurídico altamente sofisticado se funde com a competência cognitiva e emocional. Muitos dos conceitos centrais na aplicação judicial da lei como justiça, razoabilidade e intenção estão profundamente enredados no tecido da vida humana. Além disso, o raciocínio judicial combina diversas habilidades cognitivas, como avaliar fatos, interpretar textos, fazer analogias e se envolver em interações dialéticas. (SARTOR; BRANTING, 1998, tradução nossa.)

---

<sup>19</sup> Dados do Relatório Justiça em Números 2020, do. Justiça em Números 2020 ano base 2019/ Conselho Nacional de Justiça-Brasília: CNJ 2021.

<sup>20</sup> SARTOR, G.; BRANTING, L. Karl. Introduction: judicial applications of Artificial Intelligence. *Artificial Intelligence and Law*, v. 6, p. 105-110, 1998. Disponível em: [<https://bit.ly/2G6zyQz>]. Acesso em: 09 ago. 2021.

Nesse sentido, Sartor e Branting (1998) alertam que “aspectos da atividade judicial justificam uma abordagem cautelosa para os perigos de substituir a discricção judicial por um modelo rígido de computador”.

Em estudos recentes de mapeamento de aplicações já em execução no Judiciário, Chittendem<sup>21</sup> (2017) destaca uma forma de emprego da IA:

Um sistema de previsão de decisões judiciais voltado para análise de acórdãos da Corte Europeia de Direitos Humanos, o qual já atingiu índices de acertos em 79% dos casos, graças à capacidade de leitura de padrões não estritamente jurídicos, como delineamentos circunstanciais dos casos, linguagem empregada e organização dos tópicos sentenciais. (CHITTENDEM, 2017, tradução nossa).

A adoção de técnicas de aprendizado de máquina e análise de padrões para sistemas judiciários chegou ao Brasil através do projeto Victor. A ferramenta tem por finalidade auxiliar na digitalização de textos, separação e classificação de documentos do acervo e peças processuais do tribunal e identificar os temas mais destacados de repercussão geral tratados na Corte.

Em contrapartida, a ideia de solução tecnológica indicando ou apontando caminhos para a condução de litígios traz algumas preocupações sensoriais. Consoante Sartor e Branting (1998), o processo decisório de litígios envolve expertises legais, mas também competências cognitivas e emocionais. Os autores recordam que uma série de conceitos jurídicos possui valor indeterminado, “sendo completados graças a um conjunto dinâmico de percepções e sentimentos humanos de ordem íntima, leitura proativa de fatos sociais e interação dialética com a complexidade do real” (SOUSA; OLIVEIRA, 2019).

Sourdin<sup>22</sup> (2018) destaca que a atividade judiciária não se restringe a processos decisórios, mas engloba uma série de tarefas integradas.

---

<sup>21</sup> CHITTENDEM, Tara. Artificial Intelligence and the Legal Profession. Horizon Scan Forward Thinking, The Law Society, ago. 2017. Disponível em: [https://bit.ly/2TYw3QN]. Acesso em: 03 ago. 2021.

<sup>22</sup> SOURDIN, Tania. Judge v Robot? Artificial Intelligence and judicial decision-making. University of New South Wales Law Journal, v. 41, n. 4, p. 1114-1133, 2018. Disponível em: [https://bit.ly/2D2HCAP]. Acesso em: 03 ago. 2021.



A autora sintetiza quatro das principais controvérsias da aplicação de IA no Judiciário como um todo:

1. Às autoridades legais é conferida uma confiança institucional da sociedade e do Estado, o que não se estende às ferramentas tecnológicas que apontam decisões sem qualquer legitimidade prévia;
2. A transmutação de textos legais para códigos tecnológicos demandará constante atualização e ficará a cargo de programadores e outros profissionais de Tecnologia da Informação, os quais geralmente não possuem familiaridade e expertise com as fontes do sistema legal;
3. Embora as atividades jurisdicionais envolvam, por vezes, julgamentos discricionários pautados em critérios e condições muito subjetivas dos julgadores, o mero emprego de mecanismos automatizados não é capaz de retirar enviesamentos, pois estes serão alimentados por bases de dados que refletem tendências igualmente inclináveis a certos rumos; e
4. Há uma clássica diferença entre semântica e sintática, sendo que, embora a linguagem binária, natural a sistemas computacionais, permita o perfeito processamento de informações, não está apta a incorporar os significados intrínsecos atrelados a expressões e situações particulares. (SOURDIN, 2018.)

Percebe-se diversas visões em torno da IA. Mesmo quando governos e legisladores não estejam muito bem conscientes sobre qual seria a conotação mais delimitada em torno de IA, eles estão determinados a conduzir o debate e a implementação de políticas sobre o conjunto das tecnologias implicadas e seus desdobramentos sobre a vida humana.

Em outros casos, como ocorre sobretudo em países como Canadá, Índia e Reino Unido, as estratégias nacionais de IA buscam estabelecer um desenho institucional sólido para atrelar a atuação de governos, indústria, academia e sociedade civil, em modelos muito semelhantes com aqueles já

existentes da composição multissetorial na proteção do meio ambiente, direitos humanos e governança da Internet (POLIDO, 2021).

## 3. Impactos do Projeto Athos no processo decisório do STJ

### 3.1. Introdução o STJ inserido na era da inteligência artificial

No dia 14 de junho de 2018, o Superior Tribunal de Justiça deu o primeiro passo no sentido de sua inclusão na era da Inteligência Artificial, ao publicar a Instrução Normativa (IN) STJ/GP N. 6 de 12 de junho de 2018. A referida IN instituiu o projeto-piloto aplicado às soluções de IA no âmbito do STJ, com os seguintes objetivos<sup>23</sup>:

- I. avaliar a viabilidade de aplicação das soluções de Inteligência Artificial no fluxo processual da Secretaria Judiciária;
- II. propor soluções visando aumentar a produtividade e a eficácia do trabalho realizado pelas unidades;
- III. promover a melhoria do sistema classificatório dos processos e da qualidade dos dados para fins de gestão da informação e de cumprimento das diretrizes estabelecidas pelo CNJ;
- IV. contribuir para automação e racionalização das rotinas de trabalho do Tribunal;
- V. criar condições para redução do quantitativo de estagiários.

Assim, a alta administração do Tribunal constituiu uma comissão intersetorial e multidisciplinar formada por servidores da Secretaria Judiciária, Secretaria da Tecnologia da Informação e da Coordenadoria de Auditoria da Tecnologia da Informação, aos quais couberam o planejamento e coordenação das atividades, bem como a apresentação de relatórios contendo os resultados alcançados com o projeto.

Com base na composição intersetorial e multidisciplinar da Comissão, foi escolhida a metodologia CRISP-DM (*Cross Industry Standard Process for Data Mining*). Esta metodologia foi criada em 1996 (naquela época

---

<sup>23</sup> SUPERIOR TRIBUNAL DE JUSTIÇA: Relatório de encerramento projeto Athos. Rev. 06.03.

mais conhecido como mineração de dados, daí o DM em CRISP-DM). A primeira versão do CRISP-DM 1.0 foi idealizada pelas equipes da DaimlerChrysler, SPSS e NC4.

À época, a comissão de estudos, considerou o framework de processos mais utilizado pela indústria e academia<sup>24</sup>. Sua abordagem consiste em dividir um problema em 6 fases diferentes que interagem entre si de forma cíclica. As fases são:

- I. Entendimento do negócio – Especifica o problema dentro do contexto de negócio, define os objetivos e as métricas de sucesso esperadas.
- II. Entendimento dos dados – Levanta os dados existentes que possam ser utilizados para construção da solução.
- III. Preparação dos dados – Adequa e transforma os dados para a correta utilização.
- IV. Modelagem – Treina, testa, otimiza os modelos de inteligência artificial.
- V. Avaliação – Compara os diversos modelos gerados e escolhe o que melhor contribuiu para o alcance dos critérios de sucessos definidos.
- VI. Implantação – Transforma o modelo em um produto para que seja passível de utilização no fluxo de processos de trabalho da área de negócio.

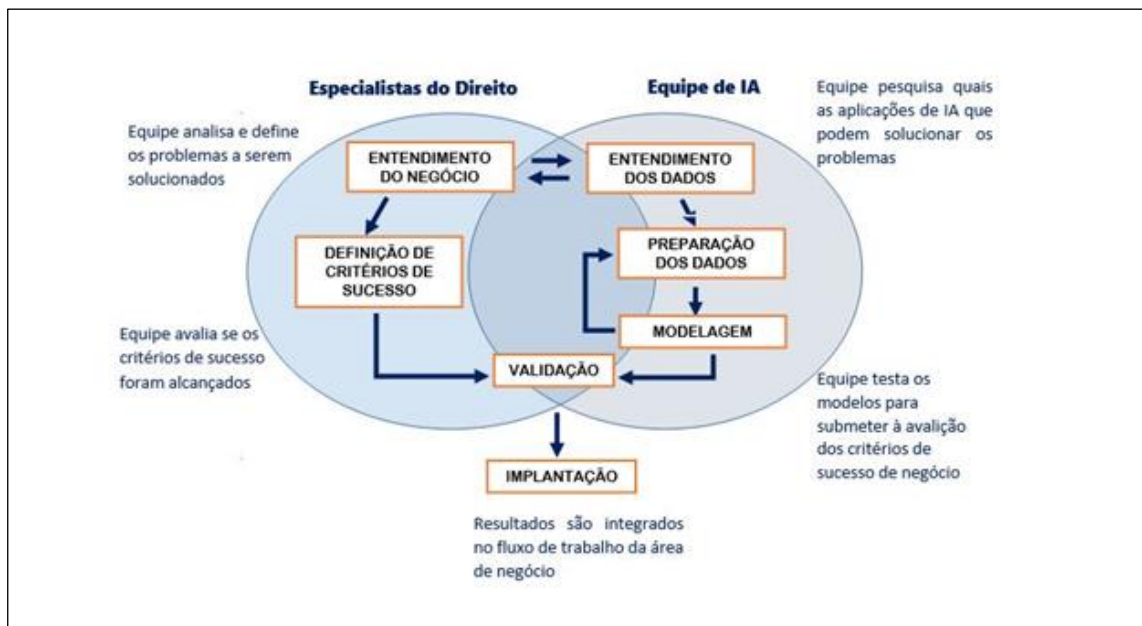
Ademais, o fluxo fornecido pelo CRISP-DM permitiu uma comunicação contínua entre os especialistas de negócio da área judiciária e os especialistas de tecnologia, dos envolvidos no projeto. A figura 2 demonstra o relacionamento entre as equipes com base nas fases do processo<sup>25</sup>.

---

<sup>24</sup> SUPERIOR TRIBUNAL DE JUSTIÇA: Relatório de encerramento projeto Athos. Rev. 06.03.

<sup>25</sup> SUPERIOR TRIBUNAL DE JUSTIÇA: Relatório de encerramento projeto Athos. Rev. 06.03.

Figura 2- Fluxo CRISP-DM.



Fonte: Relatório de encerramento Projeto Athos.

Os resultados do estudo do projeto piloto foram apresentados segmentados em três partes:

- Primeira parte: classificação automática dos processos recursais de acordo com a Tabela Unificada de Assuntos – TUA criada pela Resolução CNJ n. 46, de 18 de dezembro de 2007;(art. 2º, PU, I)
- Segunda parte: extração automática dos dispositivos legais apontados como violados pelo recorrente e indexação desses dados no sistema informatizado (indexação legislativa) para fins de triagem, a partir da análise textual da peça do recurso especial; (art. 2º, PU, II)
- Terceira parte: avaliar a viabilidade de aplicação das soluções de Inteligência Artificial no fluxo processual da Secretaria Judiciária; propor soluções visando aumentar a produtividade e a eficácia do trabalho realizado pelas unidades; (art. 2º, I e II)

A comissão responsável pelo projeto concluiu, diante dos resultados colhidos, pela viabilidade da aplicação de soluções de Inteligência Artificial no fluxo processual da Secretaria Judiciária do STJ<sup>26</sup>.

Os testes realizados revelaram que a introdução do novo padrão tecnológico apresentou grande potencial de impacto nas rotinas de trabalho avaliadas, abrindo a perspectiva de automação e racionalização de diversas tarefas na SJD, medidas que vão ao encontro das necessidades então verificadas no Superior Tribunal de Justiça.

## 3.2. Modelagem dos dados

Após os estudos realizados pela comissão intersetorial e multidisciplinar do projeto piloto, onde concluíram pela viabilidade da aplicação de soluções de IA no fluxo processual da Secretaria Judiciária do STJ, deu-se início ao projeto Athos, através de uma iniciativa desenvolvida pela Secretaria Judiciária, Núcleo de Gerenciamento de Precedentes, Secretaria de Tecnologia da Informação e Assessoria de Inteligência Artificial, daquela Corte.

O Sistema Athos, em homenagem ao ministro Athos Gusmão Carneiro, possui seu modelo de IA idealizado a partir da tese de mestrado do servidor Amilar Domingos Moreira Martins, e tem como objetivo: buscar auxiliar o agrupamento de feitos, visando a geração de decisões iguais com base em modelos precedentes. A partir da identificação deste precedente semântico, a produção da solução judicial para o caso em análise tende a ser mais rápida do que aquela formulada sem a identificação prévia. Neste contexto, a celeridade processual contribui para decisões a um custo unitário menor, prestigiando também o princípio da economicidade. (MARTINS, 2018).

O modelo de Inteligência Artificial desenvolvido no projeto Athos aplica técnicas de Inteligência Artificial no agrupamento de documentos jurídicos no âmbito do Superior Tribunal de Justiça. O objetivo foi alcançado com a produção de um modelo de inteligência artificial capaz de agrupar documentos jurídicos semanticamente semelhantes.

---

<sup>26</sup> SUPERIOR TRIBUNAL DE JUSTIÇA: Relatório de encerramento projeto Athos. Rev. 06.03.

O modelo de inteligência artificial do Athos foi treinado com o uso de acórdãos indexados pela Secretaria de Jurisprudência do STJ (SJR) entre os anos de 2015 e 2017 num total de 328.732 documentos. Os documentos foram fornecidos pela Secretaria de Tecnologia da Informação e Comunicação do STJ (STI) e seu lapso temporal determinado apenas pela disponibilidade imediata do acesso. Esse conjunto de documentos foi denominado CORPUS, que somaram um total de 643Mb, englobaram um total de 318.122 processos, e geraram um dicionário de 49.165 palavras únicas. (MARTINS, 2018).

Abaixo, vemos um exemplo do conteúdo do campo EMENTA, disponibilizado para treinamento do modelo de IA do Athos.

ADMINISTRATIVO E PROCESSUAL. AGRAVO REGIMENTAL. MANDADO DE SEGURANÇA. PROCESSO ADMINISTRATIVO DISCIPLINAR. PENA DE DEMISSÃO EM RAZÃO DE IMPROBIDADE. OPERAÇÃO EUTERPE. ALEGAÇÃO DE VIOLAÇÃO DO DEVIDO PROCESSO LEGAL. PARCIALIDADE (SUSPEIÇÃO) NÃO COMPROVADA. LEGÍTIMA UTILIZAÇÃO DA PROVA EMPRESTADA. AUSÊNCIA DE REFORMATIO IN PEJUS. INDEFERIMENTO LIMINAR MANTIDO. SEGURANÇA DENEGADA. Histórico da demanda 1. Trata-se de procedimento Administrativo Disciplinar que resultou em demissão no Ibama em razão de improbidade administrativa. Conforme consta de Relatório Final da Comissão Processante, 'Desmantelou a Polícia Federal na Operação Euterpe, o mundo criminoso instalado no âmbito do meio ambiente, atingindo o cerne da quadrilha, o núcleo interno, formado por vários segmentos de servidores do Ibama/RJ, precipuamente fiscais e técnicos e o externo, que se valia do poder financeiro para proteger seus interesses ilegais'. 2. Consta do Relatório Final da Comissão Processante a descrição das condutas investigadas que deram origem à presente impetração, que 'O investigado Leonardo mantém uma forte relação extra-funcional com (..), empresário na área de construção e de consultoria na área de meio ambiente. Na maioria das conversas existentes e degravadas neste relatório, identifica-se práticas criminais ocorridas entre ambos. (..) Verifica-se a ocorrência de fiscalização por parte do Ibama em obra do interlocutor de Maia. Inclusive nestas ligações, as conversas giram entorno de quanto, em valores, pode-se pagar. Posteriormente, Maia, em conversa com Leonardo, na data de (..) informa a este que a pessoa bateu o pé em valor de dois mil reais, que foram depositados na conta de Maia através de cheque conforme conversa (..) entre Maia e Isidoro.

Leonardo também serve de intermediário entre Maia e os outros fiscais do Ibama. Quando alguma obra é fiscalizada pelo pessoal do Ibama, os empresários ligam para Maia, este descobre quem foi o fiscal, em seguida liga para Leonardo. Efetiva o contato entre as partes (...). Leonardo também indica os serviços de Marcos Maia para empresas que ele mesmo fiscaliza, é o caso da Construtora Ontra, a qual Leonardo liga na data de (...) para Maia e informa que o pessoal da Ontra vai procurá-lo em nome dele.

Em seus estudos, Martins (2018, p. 36) identificou a necessidade de realizar alguns procedimentos de pré-processamento para ajustes e padronização do CORPUS de treinamento do modelo, pois o texto do documento trouxe uma série de redundâncias e especificidades que dificultaram sua utilização no treinamento do modelo. Abaixo, listamos os procedimentos realizados pela equipe de IA do STJ no pré-processamento do CORPUS de treinamento do Athos:

### 3.2.1. Conversão para minúsculas

Todos os caracteres dos documentos que compõem o corpus serão convertidos para minúsculo. Essa padronização garante que não haverá diferenciação entre termos iguais grafados em maiúsculas ou em minúsculas.

### 3.2.2. Remoção de *StopWords*

São consideradas *stopwords* termos que, devido ao seu uso em larga escala, não agregam informações úteis para a construção do modelo de inteligência artificial. No experimento foram removidos dos textos advérbios, conjunções, siglas de unidades da federação, algarismos romanos, preposições, pronomes, eventuais caracteres soltos e as conjugações dos verbos de ligação mais comuns (ser, estar, parecer, andar, viver, ficar, tornar, tornar-se, virar, permanecer e continuar).

### 3.2.3. Remoção de termos entre parênteses:

Da análise inicial dos documentos do corpus, verificou-se que o conteúdo entre parênteses, no maior número de vezes, não agrega conteúdo útil



ao modelo de IA. Por conta dessa observação, optou-se por remover todos os textos entre parênteses.

### 3.2.4. Remoção de Não-letas:

Todos os caracteres que não correspondem a letras foram removidos, isso inclui algarismos arábicos, hifens, parênteses, aspas, barras e traços.

### 3.2.5. Singularização de Verbetes:

Um algoritmo de conversão de plurais em singulares foi aplicado sobre o texto. Isso tende a possibilitar uma melhora no cálculo da frequência dos termos ao igualar singulares e plurais (p. ex.: réu preso = réus presos).

### 3.2.6. Conversão de caracteres especiais:

Todos os caracteres acentuados e a cedilha foram convertidos para suas versões básicas. Essa atividade garante que eventuais erros de grafia sejam ignorados na construção do modelo.

Ao final do pré-processamento, o texto do exemplo passou a apresentar a seguinte aparência:

administrativo processual agravo regimental mandado seguranga processo  
administrativo disciplinar pena demissao razao improbidade operacao  
euterpe alegacao violacao devido processo legal parcialidade comprovada  
legitima utilizacao prova emprestada ausencia reformatio peju indeferimento  
liminar mantido seguranga denegada historico demanda trata procedimento  
administrativo disciplinar resultou demissao ibama razao improbidade  
administrativa consta relatorio final comissao processante desmantelou  
policia federal operacao euterpe mundo criminoso instalado ambito meio  
ambiente atingindo cerne quadrilha nucleo interno formado vario segmento  
servidor ibama precipuamente fiscal tecnico externo valia poder financeiro  
proteger interesse ilegal consta relatorio final comissao processante  
descricao conduta investigada deram origem presente impetracao  
investigado leonardo mantem forte relacao extra funcional empresario area  
construcao consultoria area meio ambiente maioria conversa existente  
degravada relatorio identifica pratica criminal ocorrida ambos erifica  
ocorrancia fiscalizacao parte ibama obra interlocutor nesta ligacao conversa  
giram entorno valor pode pagar conversa leonardo data informa pessoa bateu  
pe valor doil mil real depositado conta atrave cheque conversa isidoro  
leonardo serve intermediario fiscal ibama alguma obra fiscalizada pessoal  
ibama empresario ligam descobre fiscal seguida liga leonardo efetiva contato  
parte leonardo indica servico empresa fiscaliza caso construtora ontra  
leonardo liga data informa pessoal ontra vai procura nome restou comprovado  
apurado acusado leonardo edward rose respeitou norma legal atuando  
acordo interesse desleal instituicao arredio atribuicao cargo resultou

demonstrado apurado forte consistente comprovacao acusado leonardo edward rose associado conduta obteve vantagem pecuniaria ilicita detrimento dignidade funcao publica utilizando condicao servidor publico precedente julgamento curso primeira julgou doil mandado seguranca semelhante Identificação de ngramas

Outro procedimento de preparação do Corpus de Treinamento do Athos é a identificação de ngramas, definida assim por Martins (2018):

Ngramas são palavras que atuam mais comumente juntas em determinado contexto do que separadas. A título de exemplo, em um contexto jurídico, a reunião dos termos “Superior”, “Tribunal” e “Justiça” tem um valor semântico diferente daqueles que os termos assumem quando separados. Os ngramas mais comuns são os bigramas, que reúnem dois termos, e os trigramas, que reúnem três. Dentro de um modelo de IA voltado para PLN, os ngramas são importantes ao garantir que termos que apresentam um valor semântico específico quando reunidos sejam tratados como um termo único. Isso contribui para a redução da dispersão do modelo. (MARTINS, 2018).

A técnica empregada, no Athos, consiste na aplicação recursiva do algoritmo de identificação de ngramas sobre o corpus de forma a que o primeiro ciclo identifique bigramas, o segundo identifique trigramas/quadrigramas o terceiro identifique quadrigramas/pentagramas e assim sucessivamente. Ao final do processamento, foram identificados 48.899 bigramas/trigramas no corpus de treinamento. Esses ngramas são tratados, durante o treinamento do modelo, como um único termo (MARTINS, 2018). Com a aplicação dos ngramas, o texto do exemplo recebeu a constituição abaixo:

Administrativo\_processual agravo\_regimental\_mandado\_seguranca processo\_administrativo\_disciplinar pena\_demissao razao improbidade operacao euterpe alegacao\_violacao devido\_processo\_legal parcialidade comprovada legitima utilizacao\_prova emprestada ausencia reformatio\_peju indeferimento\_liminar mantido seguranca\_denegada historico\_demanda trata procedimento\_administrativo\_disciplinar resultou demissao ibama razao improbidade\_administrativa consta relatorio\_final comissao\_processante desmantelou policia\_federal operacao euterpe mundo criminoso instalado ambito meio\_ambiente atingindo cerne quadrilha nucleo interno formado vario segmento servidor ibama precipuamente fiscal tecnico externo valia poder financeiro proteger interesse ilegal consta relatorio\_final comissao\_processante descricao\_conduta investigada deram origem\_presente impetracao investigado leonardo mantem forte relacao\_extra funcional empresario area construcao consultoria area meio\_ambiente maioria conversa existente degravada relatorio identifica pratica criminal ocorrida ambos erifica ocorrencia fiscalizacao parte ibama obra interlocutor nesta ligacao conversa giram entorno valor\_pode pagar conversa leonardo data informa

pessoa bateu pe valor doil mil\_real depositado\_conta atrave cheque conversa isidoro leonardo serve intermediario fiscal ibama alguma obra fiscalizada pessoal ibama empresario ligam descobre fiscal seguida liga leonardo efetiva contato parte leonardo indica servico empresa fiscaliza caso construtora ontra leonardo liga data informa pessoal ontra vai procura nome restou\_comprovado apurado acusado leonardo edward rose respeitou norma\_legal.

Martins (2018) encontrou na literatura diversas abordagens para construção de modelos de inteligência artificial voltadas ao processamento de linguagem natural. No entanto, os exemplos que apresentam maior acurácia, com objetivo semelhante ao proposto no modelo de IA do Athos, convergiram para a aplicação de vetores de palavras ou seus derivados. Por esse motivo, a opção utilizada no modelo de IA do Athos foi pelo *Paragraph Vector*.

O modelo ainda conta com o framework Gensim que é voltado ao processamento de linguagem natural. Esta ferramenta implementa a linguagem Python<sup>10</sup> nos mais populares algoritmos da área, em especial a classe Doc2Vec, que traz uma implementação do *Paragraph Vector de Le & Mikolov* (MARTINS, 2018).

A partir da organização realizada pelo procedimento de pré-processamento, passa-se ao treinamento do modelo. Neste treinamento, apresenta-se 300 mil acórdãos ao Athos, e cada documento é lido 1000 vezes por ele, possibilitando, ao sistema, a criação de um pequeno dicionário semântico. No dicionário citado, a ferramenta de IA classifica cada palavra, rotulando-as e impondo um fator crescente, de acordo com a frequência que elas aparecem nos textos, assim como o contexto em que estão inseridas, organizando os termos e documentos num espaço vetorial de 300 dimensões.

A análise semântica é definida como uma “decomposição” do corpus textual em unidades de análise “a partir da qual se torna possível uma reconstrução de significados que apresentem uma compreensão mais aprofundada da interpretação de realidade do grupo estudado”. Ou seja, a análise de conteúdo expande a capacidade de comunicação do texto ou discurso original.

A princípio, foram utilizados acórdãos recorridos e as petições de recurso especial como insumos iniciais, porém, no decorrer do projeto houve demonstrações de que o modelo se comporta adequadamente também quando

submetido à petições de habeas corpus e decisões exaradas no STJ. Isso indicou ganhos potenciais de sua utilização em outras iniciativas tais como análise de principais e sucessivos, triagem de processos em gabinetes de Ministros e triangulação de jurisprudência. Atualmente, a base do sistema ATHOS possui: 2,3 milhões de processos com 8,6 milhões de peças vetorizadas e indexadas para pesquisa semântica e textual.

### 3.3. O sistema Athos e o rito dos recursos repetitivos

Sanseverino e Marchiori (2021), no artigo “*O Projeto Athos de inteligência artificial e o impacto na formação dos precedentes qualificados no Superior Tribunal de Justiça*” esclarecem que o Código de Processo Civil de 2015 exigiu uma mudança de postura muito grande dos operadores do direito, principalmente do Poder Judiciário, pois suas disposições conduzem à conclusão de que devemos ter no Brasil menos julgamentos repetidos em demandas de massa e mais julgados paradigmas. Os autores afirmam que:

A lógica processual do CPC/2015 parece irretorquível. Deixando o Poder Judiciário de julgar a mesma questão várias vezes, ganha-se em produtividade, com a solução coletivizada de questões repetitivas, e em racionalidade, pois, para o fim da demanda repetitiva, não será mais necessária a sua tramitação em todas as instâncias. Ganha-se ainda em segurança jurídica, sendo inclusive, um importante incentivador para que novas ações não precisem ser ajuizadas no Poder Judiciário. (SANSEVERINO; MARCHIORI, 2021).

É nesse ponto que a sistemática de precedentes qualificados, consolidada pelo Código de Processo Civil de 2015, muito pode contribuir para o nosso sistema processual, deixando-o mais racional e previsível e para o direito material, pois, quando complementado por um precedente estável, íntegro e coerente, efetiva-se o ideal da necessária isonomia perante a lei e perante as decisões judiciais. (SANSEVERINO; MARCHIORI, 2021).

Há no Brasil, conforme alertam Dierle Nunes e Alexandre Bahia<sup>27</sup>, uma notória preocupação com a quantidade de julgamentos, o que

---

<sup>27</sup> BAHIA, Alexandre Gustavo Melo Franco; NUNES. Dierle José Coelho. Processo constitucional: uma abordagem a partir dos desafios do Estado Democrático de Direito. Revista Eletrônica de Direito Processual, v. 4, n. 4, 2009.

representa eficiência quantitativa, mensurada por índice de produtividade. Consequentemente, no STJ, a Presidência e os Ministros estão atentos a essa questão, cientes de que a estratégia solitária de apostar somente na ampliação quantitativa do volume de julgamentos certamente não conseguirá efetivamente atingir o ideal de produtividade e, principalmente, de racionalidade do sistema processual brasileiro. (SANSEVERINO; MARCHIORI, 2021).

O Ministro do STJ, Gurgel de Faria, alerta que é alto o percentual de processos que não mereceriam ser apreciados pelo Superior Tribunal de Justiça, chegando em até 70% do que é distribuído aos Ministros. O magistrado afirma que “precisamos ter um número maior de temas; “até agora são mil, mas há quanto tempo estamos julgando repetitivos?”. O uso de ferramentas de inteligência artificial é um dos instrumentos para a consecução desse objetivo maior de formação e julgamento de precedentes qualificados no STJ, sendo uma exigência inseparável da tecnologia a mudança da forma de atuar de todo o Poder Judiciário<sup>28</sup>.

O Superior Tribunal de Justiça recebe em média mil e quinhentos processos por dia útil, sendo natural que, nesse quantitativo, existem inúmeros processos idênticos ou semelhantes, constituindo um dos principais desafios históricos da sua Secretaria Judiciária a identificação das correlações entre os processos (SANSEVERINO E MARCHIORI, 2021). Portanto, na recorribilidade extraordinária (ou especial) o STJ recebe uma infinidade de recursos dos trinta e dois tribunais a ele vinculados em formato ainda digitalizado

De acordo com o Painel de Consulta ao Banco de Dados Nacional de Demandas Repetitivas e Precedentes Obrigatórios do Conselho Nacional de Justiça<sup>29</sup>, 1.213.188 processos deixaram de ingressar no STJ, ao se enquadrarem em temas repetitivos cadastrados na base de dados daquela Corte Superior, foram julgados no tribunal de origem. Somente a Justiça Estadual foi

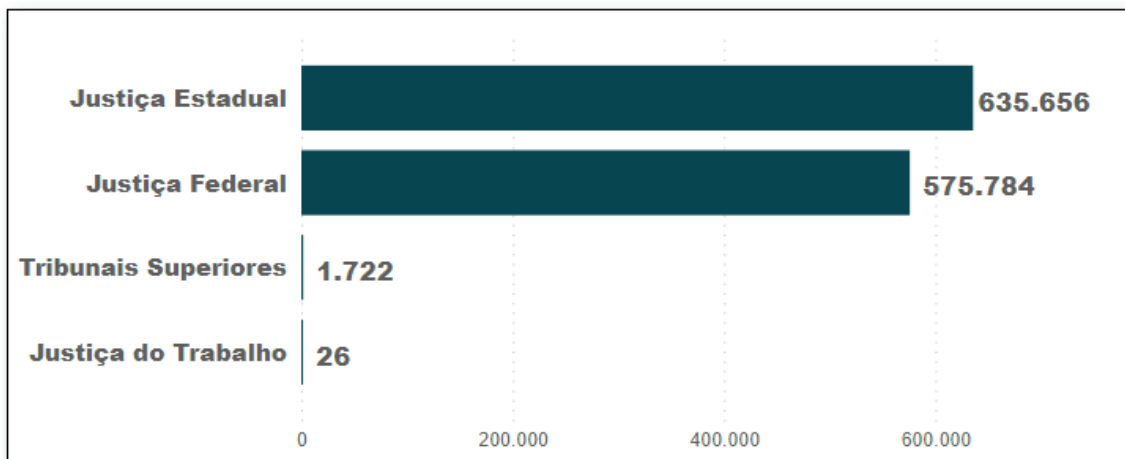
---

<sup>28</sup> Sanseverino, Paulo de Tarso; Marchiori, Marcelo Ornellas: O Projeto Athos de Inteligência Artificial e o impacto da formação dos precedentes qualificados no Superior Tribunal de Justiça.

<sup>29</sup> CONSELHO NACIONAL DE JUSTIÇA: Painel de Consulta ao Banco de Dados Nacional de Demandas Repetitivas e Precedentes Obrigatórios. Disponível em: <[https://paineis.cnj.jus.br/QvAJAXZfc/opendoc.htm?document=qvw\\_/PaineICNJ.qvw&host=QVS@neodimio03&anonymos=true&sheet=shDRGraficos](https://paineis.cnj.jus.br/QvAJAXZfc/opendoc.htm?document=qvw_/PaineICNJ.qvw&host=QVS@neodimio03&anonymos=true&sheet=shDRGraficos)>. Acesso em: 03 fev. 2022.

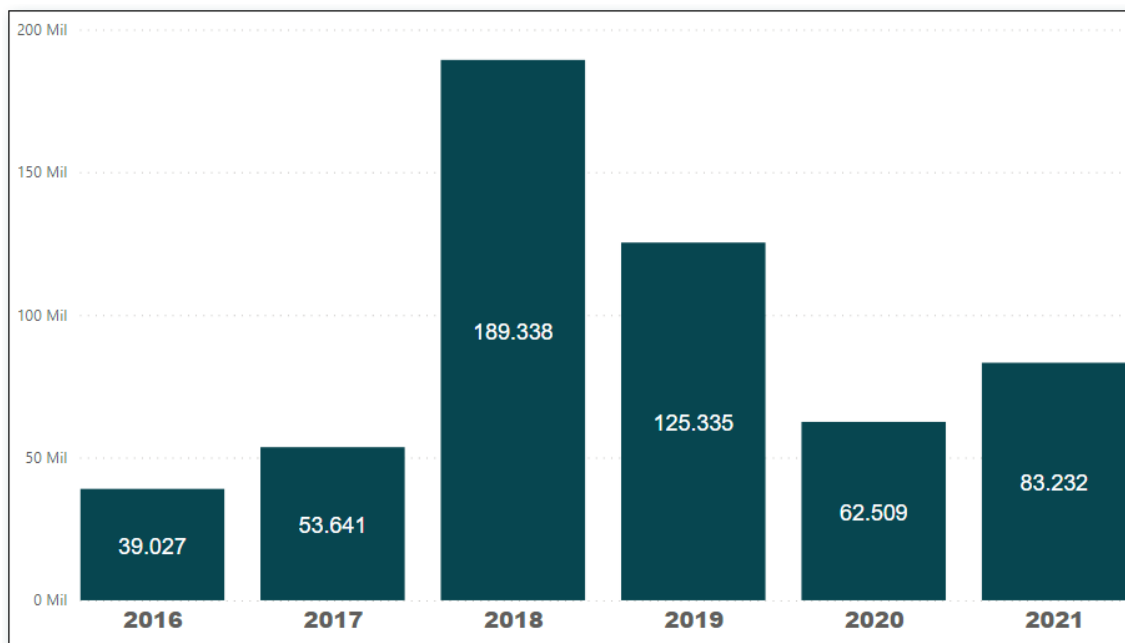
responsável pelo julgamento de 635.656 processos, logo em seguida vem a Justiça Federal com 575.784 processos e os Tribunais Superiores com 1.722 e Justiça do Trabalho com 26, como visto no Gráfico 2.

Gráfico 2- Processos julgados a partir de temas do rito dos repetitivos



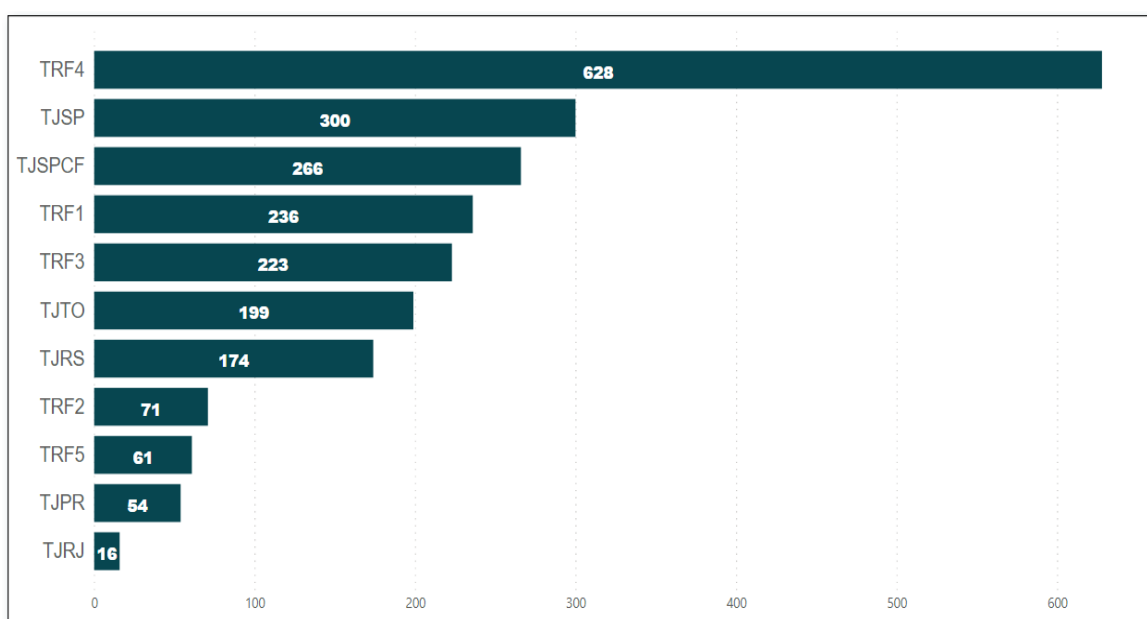
O Gráfico 3 ilustra a série histórica dos últimos 6 (seis) anos, com a quantidade de processos julgados em todo Poder Judiciário a partir dos temas repetitivos definidos pelo Superior Tribunal de Justiça:

Gráfico 3- Série histórica processos julgados RR-STJ



Ademais, o STJ possui o Núcleo de Admissibilidade e Recursos Repetitivos (NARER), responsável pela triagem dos processos que ingressam no Tribunal e que possam ser submetidos à afetação para julgamento sob o rito dos recursos repetitivos. Somente no ano de 2021, o NARER identificou 1.422 processos que foram afetados e julgados sob o rito dos RR's. Dados disponibilizados pela Secretaria Judiciária do STJ (SJD), demonstram os tribunais com maiores números de processos enviados ao STJ, em 2021, mesmo com temas já definidos e cadastrados pela Corte Superior, o Gráfico 4 ilustra esta informação:

Gráfico 4-Tribunais que enviam processos com temas definidos



De acordo com os artigos 46-A e 256 e seguintes do Regimento Interno em conjunto com a delegação da Presidência do STJ contida na Portaria STJ/GP n. 299 de 19 de julho de 201725, o Presidente da Comissão Gestora de Precedentes possui a competência de despachar em processos, antes da distribuição, com indicação para afetação para julgamento sob o rito dos repetitivos.

A Comissão Gestora de Precedentes e de Ações Coletivas – COGEPAC é uma comissão permanente de Ministros formada por três ministros do STJ, como membros efetivos e um ministro como suplente, um juiz supervisor

e três servidores assessores do Núcleo de Gerenciamento de Precedentes e de Ações Coletivas (NUGEPNAC)<sup>30</sup>.

Sua missão é coordenar os procedimentos administrativos decorrentes do julgamento de casos repetitivos, de incidentes de assunção de competência, suspensão em incidente de resolução de demandas repetitivas, bem como do monitoramento e sistematização das informações relativas ao julgamento das ações coletivas, no âmbito do Tribunal da Cidadania. A Comissão tem atuação estratégica no estabelecimento de metas e na busca das melhores soluções para o êxito do sistema de precedentes estabelecido pelo Código de Processo Civil de 2015, regulamentado internamente pelas Emendas Regimentais n. 22, 24 e 26 do Superior Tribunal de Justiça.

Sanseverino e Marchiori (2021) mapearam o fluxo de andamento dos repetitivos no STJ. Onde, inicia-se por uma triagem processual antes da distribuição de processos aos gabinetes dos Ministros, privilegiando o julgamento de mérito da questão jurídica e não a atuação individualizada da Corte. A prática consiste na identificação, pelo sistema Athos, de grupos de processos idênticos ou semelhantes entre o grande volume de processos que chegam diariamente ao STJ.

A sinalização, com a diligente assessoria dos servidores integrantes do Núcleo de Gerenciamento de Precedentes (NUGEPNAC) e da Secretaria de Jurisprudência do STJ, de processos semelhantes ou idênticos, dispara atividades devidamente mapeadas e coordenadas com a Secretaria Judiciária para que representativos daqueles processos sejam desviados do seu fluxo de distribuição, com o encaminhamento para o Gabinete do Presidente da Comissão Gestora de Precedentes. (SANSEVERINO; MARCHIORI, 2021).

Depois de uma rigorosa análise, o presidente da comissão delimita, mediante despacho fundamentado, a questão jurídica a ser submetida para julgamento sob o rito dos repetitivos nos processos e os encaminha, inicialmente, à Procuradoria-Geral da República para manifestação do *parquet*

---

<sup>30</sup> SUPERIOR TRIBUNAL DE JUSTIÇA: Disponível em: <<https://www.stj.jus.br/sites/portalp/Precedentes/informacoes-gerais/comissao-gestora-de-precedentes-e-de-acoes-coletivas>>. Acesso em: 03 dez. 2021



somente quanto à análise de o recurso tramitar na Corte Superior como uma sugestão de afetação ao rito dos repetitivos.

Após o retorno dos autos do Ministério Público, já devidamente sinalizados eletronicamente como um possível recurso repetitivo, há nova análise pelo presidente da comissão gestora de precedentes que, mediante novo despacho, lança informações importantes sobre a tese jurídica, em especial o impacto aproximado no STJ, eventual uniformidade ou diversidade de entendimentos nos tribunais e no STJ. Dessa forma, o primeiro processo é distribuído, em regra, de forma livre, sendo os demais processos conexos, representativos da mesma controvérsia, encaminhados ao relator por prevenção.

A partir do momento da conclusão do último processo, o relator possui o prazo de sessenta dias úteis para submeter a questão ao Plenário Virtual de afetação ou para deliberar, monocraticamente, os motivos pelos quais não submeterá os processos aos demais membros da Corte Especial ou Seção correspondente. Ultrapassado esse prazo de sessenta dias úteis, há uma presunção relativa de que o Ministro rejeitou a indicação do recurso como representativo da controvérsia.

A decisão mais comum, contudo, tem sido a submissão dos processos pelo relator para a análise do órgão julgador competente (Corte Especial ou respectiva Seção) para deliberar acerca da afetação do recurso ao rito dos repetitivos em ambiente eletrônico. Acolhida a proposta, é criado novo Tema repetitivo no STJ, com todos os seus efeitos jurídicos. (SANSEVERINO; MARCHIORI, 2021).

O sistema Athos possui papel estratégico fundamental no rito dos repetitivos no âmbito do Superior Tribunal de Justiça. Nesse sentido, ele é utilizado em duas vertentes distintas naquela Corte. A primeira no apoio à identificação de possíveis novos paradigmas para estabelecimento de um novo tema repetitivo, através de agrupamentos de processos idênticos ou semelhantes. A outra vertente encontra-se na Secretaria Judiciária do STJ, onde a partir da mesma metodologia de agrupamento de processos idênticos ou semelhantes o sistema Athos identifica, entre os processos que ingressam no STJ, possíveis processos que possuam tese idêntica à temas afetados e que

ainda aguardam julgamento, ou mesmo processos que já possuem temas julgados e definidos, porém o tribunal *a quo* insiste em enviá-los ao STJ.

Assim, apresentaremos os resultados do Sistema Athos, em sua primeira vertente: no apoio à identificação de possíveis novos paradigmas para estabelecimento de um novo tema repetitivo, através de agrupamentos de processos idênticos ou semelhantes. Para tanto, utilizamos dados de indicadores disponibilizados pela equipe do NUGEEPAC e log de acesso às funcionalidades do sistema, armazenados nos servidores da Assessoria de Inteligência Artificial do STJ.

## 4. As Demandas do Poder Judiciário e Os Resultados do Sistema Athos

### 4.1 A crise de demandas no poder judiciário e os recursos repetitivos

Desde os anos 30, já se discutia a sobrecarga processual do Supremo Tribunal Federal (STF). A criação do Tribunal Federal de Recursos (TFR) na década de 40, em certo sentido, foi uma resposta a essa situação. A partir dos anos 60, acentuou-se o debate, no que passou a ser chamada como “Crise do Supremo”<sup>31</sup>.

A crise consistia no “desequilíbrio entre o número de feitos protocolados e o de julgamentos por ele proferidos, sendo a entrada daqueles consideravelmente superior à capacidade de sua decisão, se acumulando os processos não julgados, resultando daí o congestionamento.

A lentidão para a efetiva prestação da tutela jurisdicional é questão que se arrasta desde a criação do STF: Alfredo Buzaid denominou como crise do STF, o descompasso entre o número de feitos protocolados e a capacidade de julgamento pela corte suprema, de sorte que, passadas algumas décadas após sua criação em 1890, constatou-se a impossibilidade material de dar vazão a julgamentos na mesma proporção que recursos e ações de sua competência davam entrada. (MANCUSO, 2018).

A crise em questão nasceu do elevado número de processos que anualmente davam entrada no Protocolo da Corte, principalmente os recursos extraordinários e seus recursos conexos (agravo de instrumento, arguições de relevância, e embargos de declaração e divergência). Em 1891 deu entrada no STF o Recurso Extraordinário n. 1; em 1904, a média anual deu em 26; em 1933, montou a média em 55; já em 1960, ano em que o STF transferiu-se para Brasília, foram julgados 5.946 recursos extraordinários, quantidade que já à época foi considerada excessiva e preocupante por seus juízes. O crescimento avantajou-se quase em “progressão geométrica”, sendo que o

---

31 SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: <https://www.stj.jus.br/sites/portalp/Institucional/Historia/A-crise-do-Supremo>. Acesso em: 08 set. 2021

relatório anual de 1985 informava terem sido julgados 17.798 processos diversos (MANCUSO, 2018).

Números posteriores davam conta de que em 1994 foram julgados 16.250 recursos extraordinários e 16.758 recursos especiais. Reportando-se aos números para 1997, de janeiro a julho daquele ano, o STF recebeu 19.599 processos e julgou 21.317. o volume de ações protocoladas no Supremo Tribunal Federal no decorrer do ano [de 2000] cresceu 49,43% em relação a 1999 – no total foram 101.996 processos contra 68.255 (MANCUSO, 2018).

Atualmente tramitam quase 77 milhões de processos em todas as instâncias e setores do Judiciário<sup>32</sup>. Só no Supremo, para se ter ideia da carga de trabalho de seus 11 ministros, até julho de 2021 foram protocolados 14.911 processos na corte, uma média de 1355 processos por ministro<sup>33</sup>.

À semelhança do que se passa no STF, também no STJ é preocupante o aumento vertiginoso de processos. No ano 2000, o Min. Costa Leite, dava a notícia da realização de um mutirão num fim de semana, em Brasília, mobilizando cerca de 100 servidores, para autuar e distribuir mais de 26 mil processos parados na corte.

Em média, o tribunal recebe atualmente 3 mil processos por semana – ou seja, mais de 10 mil por mês. A nossa capacidade de julgamento chegou ao limite. (...) Só no ano passado [1999] os 33 ministros julgaram 128 mil processos. Este ano [2000], em apenas três meses, o tribunal já recebeu mais de 26 mil ações, o que representa um aumento de 4,5%. (MANCUSO, 2018).

Atualmente, até o mês de dezembro de 2021, o STJ recebeu 408.770 processos, uma média de 34.064 processos por mês<sup>34</sup>, ou seja, 250% a mais do que o Ministro Costa Leite considerou ser o limite do tribunal há 21 anos. O gráfico 5 ilustra a quantidade de processo recebidos pelo STJ nos últimos 6 anos.

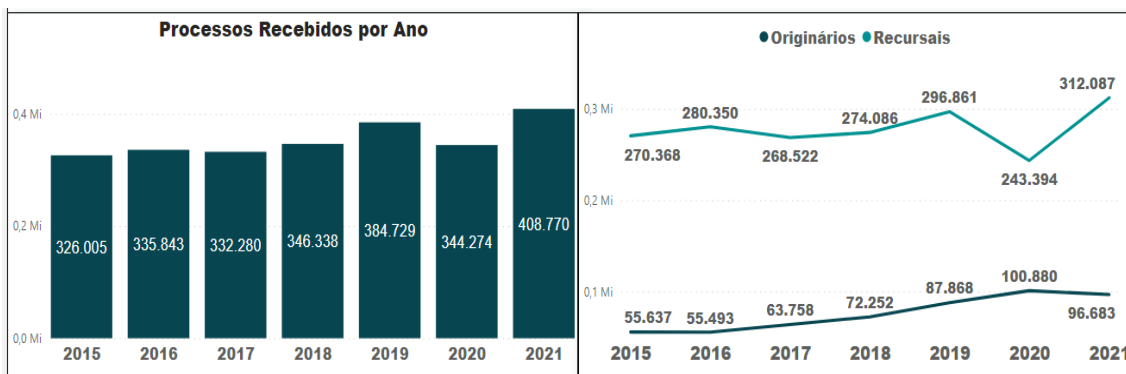
---

<sup>32</sup> CONSELHO NACIONAL DE JUSTIÇA. Justiça em Números: Disponível em: < <https://www.cnj.jus.br/pesquisas-judiciarias/justica-em-numeros/>>. Acesso em: 08 set. 2021.

<sup>33</sup> SUPREMO TRIBUNAL FEDERAL. Disponível em: < <https://bityli.com/Tuxy1>>. Acesso em: 08 set 2021.

<sup>34</sup> SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: <https://www.stj.jus.br/sites/portaip/Inicio>. Acesso em 20 jan. 2022.

Gráfico 5- Processos ingressados anualmente no STJ

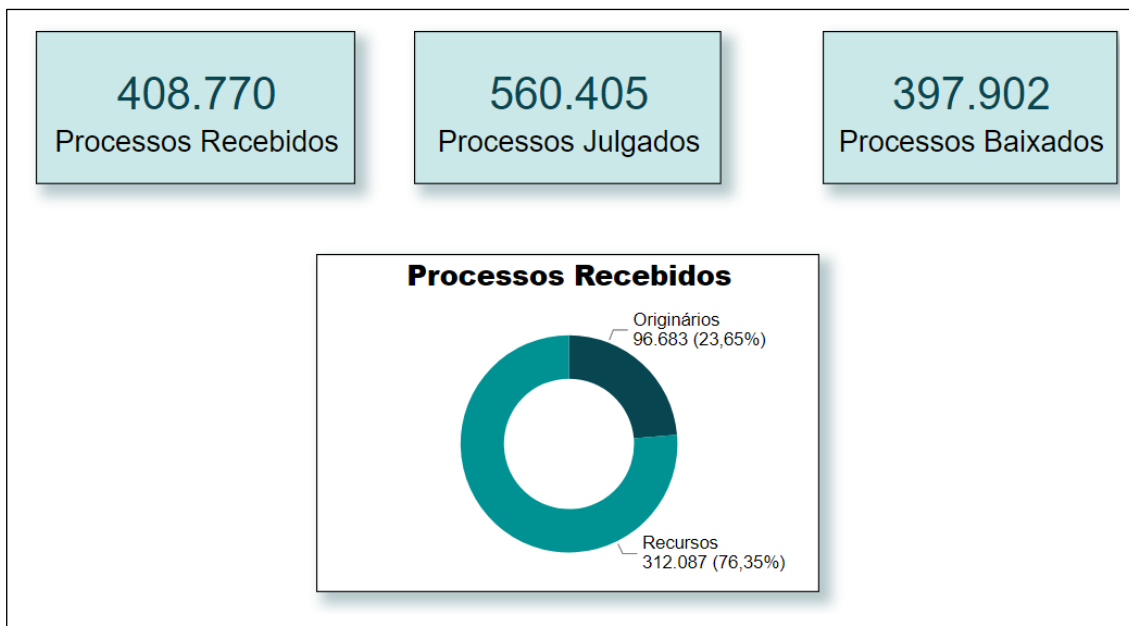


O Superior Tribunal de Justiça apresenta 99,6% do seu acervo digital, e 97% dos processos chegam de forma eletrônica ao tribunal. No mês de dezembro, o STJ recebeu 24.486 processos novos originários e recursais, totalizando 408.770 casos novos no ano. Houve aumento de 18,73% (64.496) em comparação com 2020, atingindo recorde histórico, após queda de 10,52% (40.455) no ano anterior. Fazendo uma análise mais detalhada por origem dos processos, é possível perceber que os processos recursais são os principais responsáveis por essa oscilação, com aumento de 68.693 (28,22%) em 2021, após queda acentuada de 53.467 (18,01%) em 2020. Os processos originários se mantiveram estáveis em relação à 2020 com leve queda de 4.197 (4,16%) processos, após quatro anos de crescimento contínuo entre 2016 e 2020 com média de crescimento de 16,12% ao ano no período.

O Gráfico 6 ilustra a atividade jurisdicional do STJ entre janeiro e dezembro de 2021<sup>35</sup>:

<sup>35</sup> SUPERIOR TRIBUNAL DE JUSTIÇA. Boletim Estatístico de dezembro de 2021. Disponível em: <<https://www.stj.jus.br/webstj/Processo/Boletim/sumario.asp>>. Acesso em: 15 jan. 2022.

Gráfico 6- Atividade jurisdicional do STJ entre janeiro e dezembro de 2021



A necessidade de gestão dos processos, e a racionalidade dos julgamentos, há algum tempo é pauta de preocupação no Poder Judiciário brasileiro. Paralelamente à questão do asoerbamento do Judiciário são criados mecanismos que tentam mitigar tal cenário.

A instituição de mecanismos de acesso à justiça, o aumento da litigiosidade e a busca por maior celeridade no julgamento de processos provocou modificações no ordenamento jurídico capazes de dar respostas às demandas de celeridade e segurança jurídica nos julgamentos. Uma das respostas foi a implementação da sistemática de recursos repetitivos, ainda na vigência do Código de Processo Civil de 1973, por meio da lei 11.418/2006. Essa lei inseriu o artigo 543-B no CPC, passando a regulamentar o julgamento de recursos repetitivos no âmbito do Supremo Tribunal Federal. Posteriormente, foi promulgada a lei 11.672/2008, que ficou conhecida como a Lei dos Recursos Repetitivos, estabelecendo a referida sistemática no âmbito do Superior Tribunal de Justiça <sup>36</sup>.

O objetivo dos recursos repetitivos é dar, além de mais celeridade, isonomia e segurança jurídica no julgamento de recursos especiais

<sup>36</sup> COELHO, Marcus Vinicius Furtado: Arts. 1.036 a 1.041 do CPC - Recursos repetitivos. Disponível em: <https://www.migalhas.com.br/coluna/cpc-marcado/307178/arts--1-036-a-1-041-do-cpc---recursos-repetitivos>. Acesso em: 02 set. 2021.

que tratem da mesma controvérsia jurídica. Esses casos podem ser selecionados por amostragem, onde cabe ao presidente ou vice-presidente do tribunal de origem admitir um ou mais recursos que melhor representem a questão repetitiva e encaminhá-los ao STJ para julgamento<sup>37</sup>.

O recurso representativo de controvérsia (RRC) é o processo escolhido dentre vários outros que possuam a mesma questão de direito, e que servirá como caso concreto paradigma para que o Superior Tribunal de Justiça fixe a tese jurídica, tornando-a tema repetitivo.

Por esta razão, a lei e a prática forense exigem que sua escolha seja criteriosa e pautada nos requisitos da existência de multiplicidade de recursos com idêntica questão de direito e a escolha de mais de um processo, com a maior gama de fundamentos e argumentos, para tramitação conjunta no STJ.

1. De acordo com o STJ, os recursos representativos de controvérsia poderão advir de quatro situações: Recursos admitidos pelos tribunais de origem com fundamento no art. 1.036, § 1º do CPC de 2015. Ao admitir RRCs, o tribunal de origem inicia o controle para sobrestamento de demais processos no estado ou na região em que discutida a mesma matéria.
2. Recursos admitidos pelos tribunais de origem com fundamento no art. 1.041 do CPC que, tratam de matéria que já foi firmada como tese repetitiva por tribunal superior, e não houve sua aplicação pelo tribunal de origem, em juízo de retratação, considerando-se possível distinção ou superação de tese. Ao admitir RRCs, o tribunal de origem inicia o controle para sobrestamento de demais processos no estado ou na região em que discutida a mesma matéria.
3. Recursos encaminhados ao STJ que não foram selecionadas como representativos da controvérsia pelos tribunais de origem, mas que, por meio do desenvolvimento de trabalho de inteligência no STJ, antes da distribuição, foram identificados por conterem matérias com “potencial de repetitividade ou com relevante questão de direito, de grande repercussão social, aptas a serem

---

<sup>37</sup> SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: <<https://www.stj.jus.br/sites/portalp/Precedentes/informacoes-gerais/recursos-repetitivos>>. Acesso em: 06 set. 2021.

submetidas ao Superior Tribunal de Justiça, sob a sistemática dos recursos repetitivos” (art. 46-A, IV, do RISTJ).

4. Recursos especiais interpostos contra o julgamento de mérito de um Incidente de Resolução de Demandas Repetitivas (IRDR) pelos tribunais de segunda instância (art. 256-H, do RISTJ). Isso porque a tese jurídica firmada pelo STJ, ao apreciar o mérito desse recurso, terá aplicação nacional mesmo que não seja objeto de afetação.

Todos os Recursos Representativos de Controvérsias remetidos ao STJ pelos tribunais de origem, ou selecionados pelo próprio STJ, são analisados, quanto a aspectos formais, pelo Presidente da Comissão Gestora de Precedentes e de Ações Coletivas, nos termos dos art. 256-B a 256-D e da Portaria STJ/GP n. 98/2021, o qual, por meio de despacho nos autos<sup>38</sup>:

- i. destaca a questão de direito delimitada pelo tribunal de origem indicada como repetitiva no estado ou na região, com o registro de informações importantes do processo;
- ii. determina a abertura de vista ao Ministério Público Federal para manifestação a respeito dos pressupostos de admissibilidade do recurso especial como representativo da controvérsia; determina a distribuição do recurso a um dos ministros do STJ.

Em casos de multiplicidade de recursos especiais com fundamento em idêntica questão de direito, o STJ seleciona um ou mais recursos como paradigma (ou os recebe como indicação dos tribunais de origem) para que a decisão seja replicada nas demais instâncias. Os temas são cadastrados em ordem sequencial numérica e podem ser acompanhados na página de repetitivos do tribunal.

O Banco Nacional de Dados de Demandas Repetitivas e Precedentes Obrigatórios (BNPR), foi criado pela Resolução nº 235/2016 do Conselho Nacional de Justiça (CNJ). Trata-se de uma ferramenta que possibilita consulta às informações de Incidentes de Resolução de Demandas Repetitivas (IRDR) e Incidentes de Assunção de Competência (IAC). O BNDR informa que

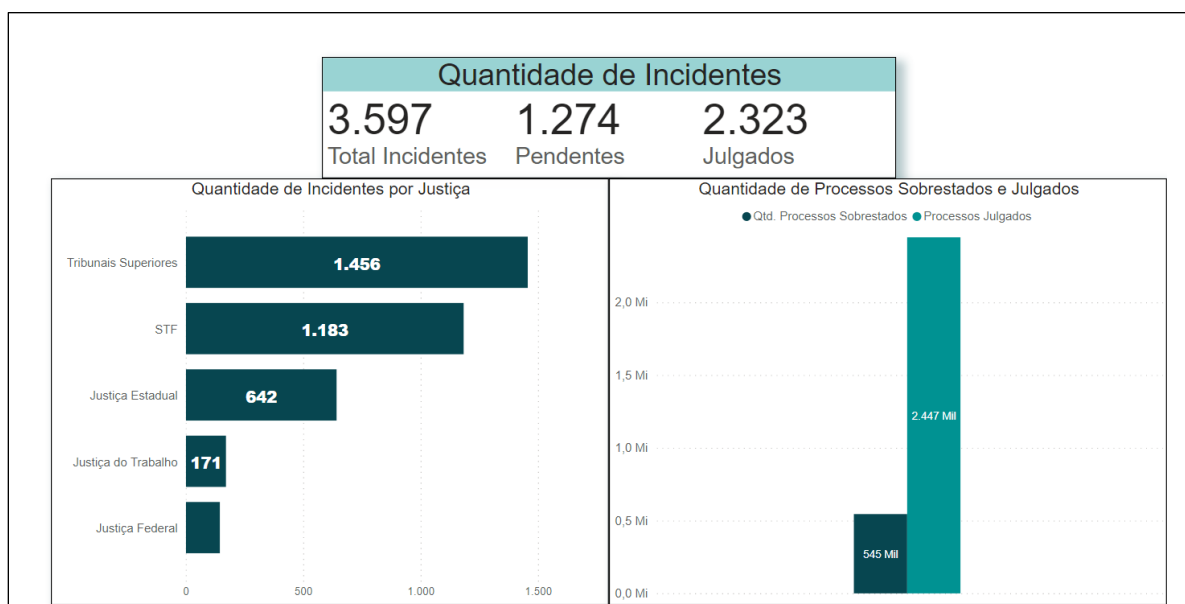
---

<sup>38</sup> SUPERIOR TRIBUNAL DE JUSTIÇA: Disponível em: <https://www.stj.jus.br/sites/portalp/Precedentes/informacoes-gerais/controversias>. Acesso em: 20 dez. 2021.



atualmente constam 3.597 incidentes de demandas repetitivas catalogados no Judiciário Brasileiro, nos quais 1.274 encontram-se pendentes e 2.323 foram julgados<sup>39</sup>. Os 3.445 incidentes catalogados impactam em 1.971.535 processos sobrestados. Atualmente, constam 2.326.184 processos julgados no BNPR. Os Tribunais Superiores concentram a maior parte dos incidentes registrados (1.456), seguido do Supremo Tribunal Federal, com 1.183 incidentes. A Justiça Estadual vem em seguida com 642, e logo atrás Justiça Federal com 145 e Justiça do Trabalho com 171 incidentes. O Gráfico 7 ilustra a distribuição destes incidentes no poder judiciário:

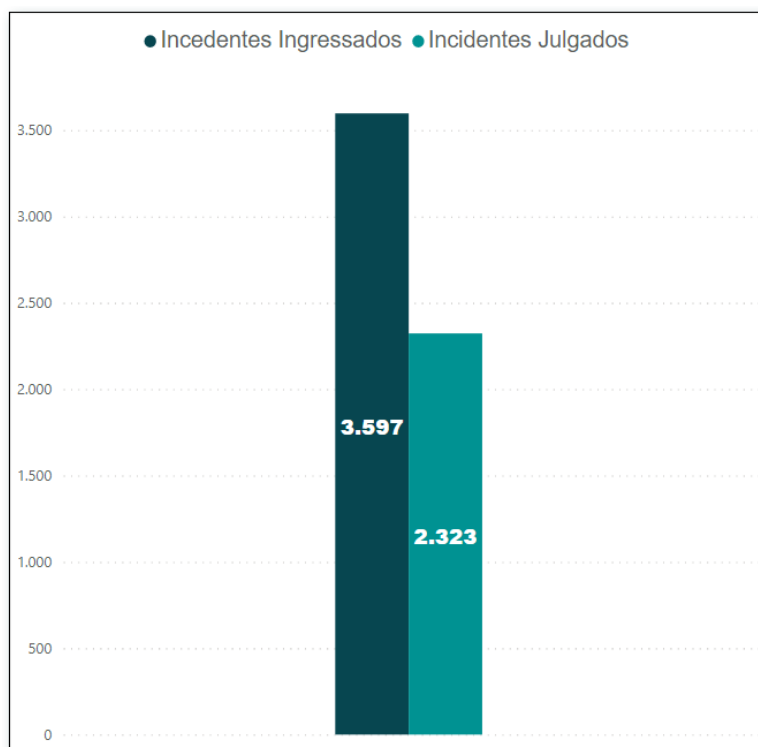
Gráfico 7- Distribuição de incidentes no poder judiciário



Observa-se também a disparidade no número de incidentes ingressados anualmente (3.597) e o número de incidentes julgados no mesmo período (2.323). Percebe-se um déficit de 35,42% na capacidade de julgamento dos incidentes ingressados anualmente no Poder Judiciário brasileiro:

<sup>39</sup> CONSELHO NACIONAL DE JUSTIÇA: Painel de Consulta ao Banco de Dados Nacional de Demandas Repetitivas e Precedentes Obrigatórios. Disponível em: <[https://paineis.cnj.jus.br/QvAJAXZfc/opendoc.htm?document=qvw\\_/PaineICNJ.qvw&host=QV S@neodimio03&anonymous=true &sheet=shDRGraficos](https://paineis.cnj.jus.br/QvAJAXZfc/opendoc.htm?document=qvw_/PaineICNJ.qvw&host=QV S@neodimio03&anonymous=true &sheet=shDRGraficos)>. Acesso em: 20 jan. 2022.

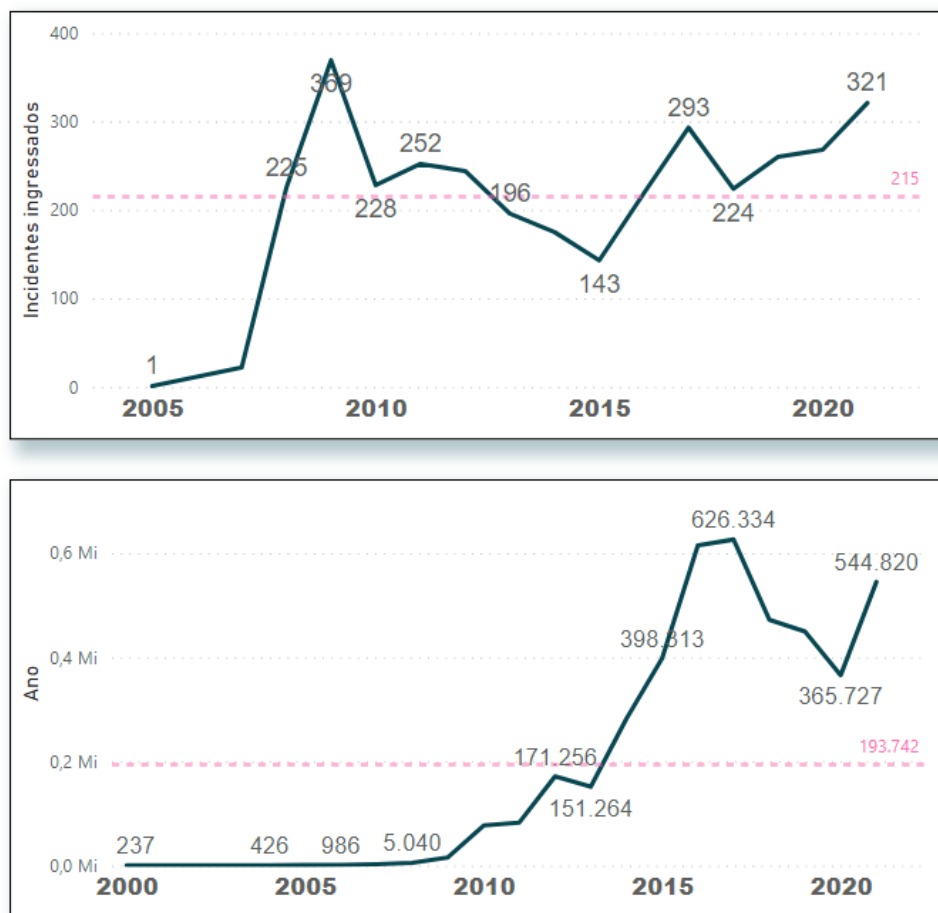
Gráfico 8- Número de incidentes ingressados anualmente/número de incidentes julgados.



Outra relação importante observada no BNDR foi entre o número de incidentes ingressados e o de processos sobrestados por ano. Percebe-se um salto no número de incidentes no ano de 2009 (369). Porém, nos anos seguintes manteve-se uma média de 215 incidentes anuais. No entanto, o número de processos sobrestados, impactados por esses incidentes, a partir de 2013, teve um salto de 306,57% entre os anos de 2013 e 2016, mantendo a média de 193.742 processos sobrestados em todo período<sup>40</sup>.

<sup>40</sup> CONSELHO NACIONAL DE JUSTIÇA: Painel de Consulta ao Banco de Dados Nacional de Demandas Repetitivas e Precedentes Obrigatórios. Disponível em: <[https://paineis.cnj.jus.br/QvAJAXZfc/opendoc.htm?document=qvw\\_/PaineICNJ.qvw&host=QVS@neodimio03&anonymous=true&sheet=shDRGraficos](https://paineis.cnj.jus.br/QvAJAXZfc/opendoc.htm?document=qvw_/PaineICNJ.qvw&host=QVS@neodimio03&anonymous=true&sheet=shDRGraficos)>. Acesso em: 20 jan. 2022.

Gráfico 9- Número de incidentes ingressados anualmente/número de processos sobrestados.



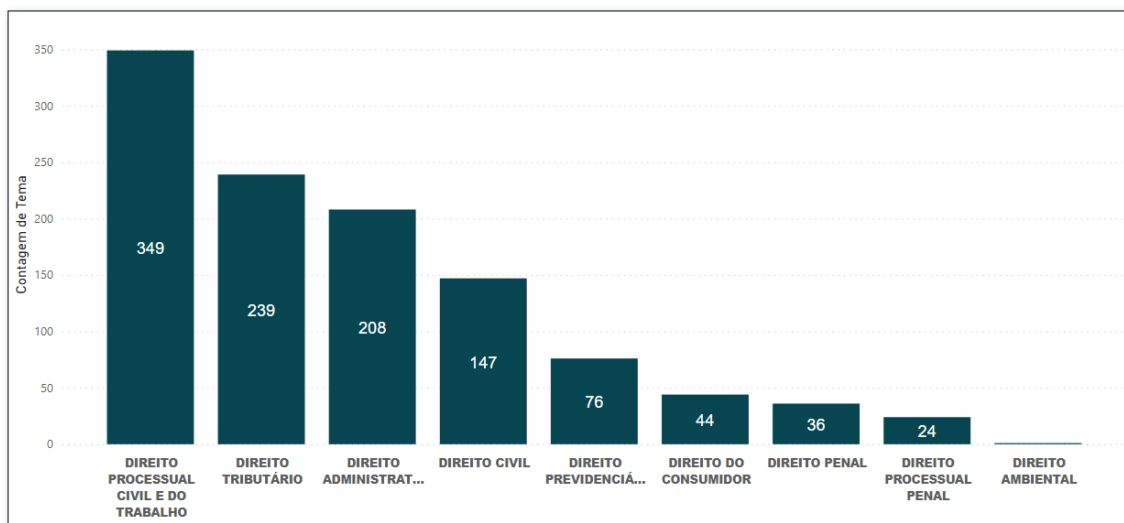
No Superior Tribunal de Justiça, o ano de 2018 representou dois marcos importantes para o sistema dos precedentes judiciais e, em especial, para a sistemática dos recursos especiais repetitivos: em agosto daquele ano, a Lei dos Recursos Repetitivos, que estabeleceu o procedimento para o julgamento de demandas de massa no STJ, completou dez anos; e no mês de novembro daquele ano, ao decidir pela afetação do Recurso Especial (REsp) 1.763.462, o STJ chegou à marca de mil temas cadastrados no sistema de repetitivos<sup>41</sup>.

O STJ possui 1.124 temas cadastrados em sua base de dados de precedentes qualificados, divididos em 9 ramos do Direito: Direito Processual Civil e do Trabalho (349), Direito Tributário (239), Direito Administrativo (208),

<sup>41</sup> SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: <https://bitly.com/cq0pg>. Acesso em: 10 set. 2021

Direito Civil (147), Direito Previdenciário (76), Direito do Consumidor (44), Direito Penal (36), Direito Processual Penal (24) e Direito Ambiental (1)<sup>42</sup>.

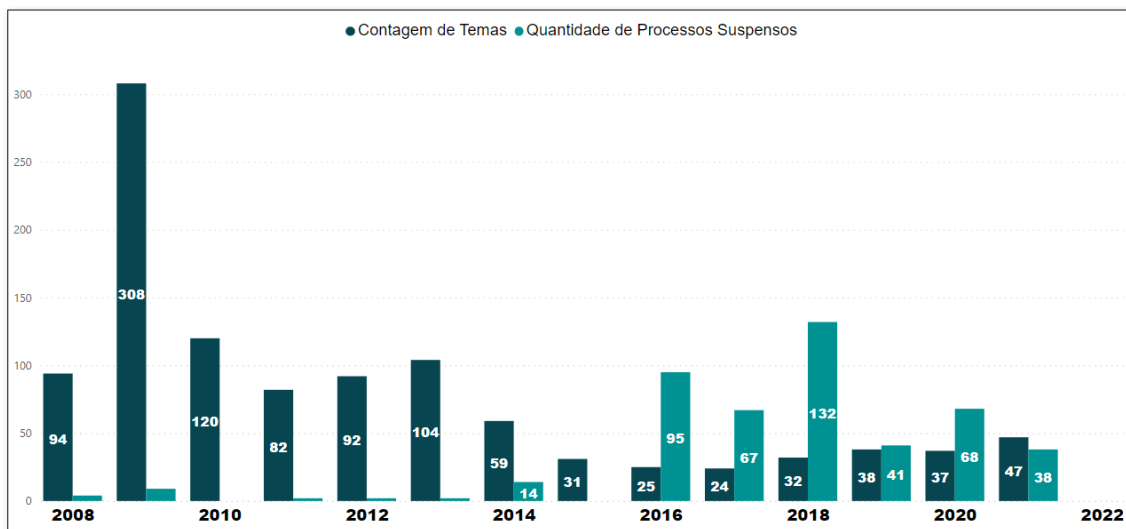
Gráfico 10- Ramos do Direito dos temas cadastrados na base de repetitivos do STJ.



Do momento da promulgação da lei 11.672/2008, no tema 1, ao tema 1.124, com a exceção do ano de 2009, onde houve uma alta de 227,66% (214) de temas cadastrados, este número vem sofrendo uma redução contínua. No entanto, de forma inversa, o número de processos suspensos, no STJ, aguardando definição da tese jurídica vem sofrendo relevante alta nos últimos 05 anos. Como demonstrado no Gráfico 11.

<sup>42</sup> SUPERIOR TRIBUNAL DE JUSTIÇA Disponível em: <[https://processo.stj.jus.br/repetitivos/temas\\_repetitivos/](https://processo.stj.jus.br/repetitivos/temas_repetitivos/)>. Acesso em 20 jan. 2022.

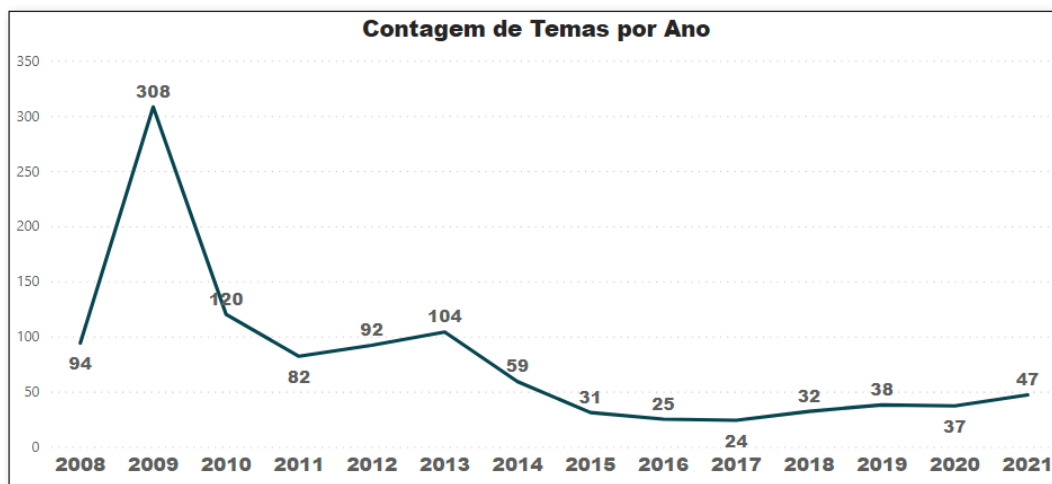
Gráfico 11- Contagem de Temas/Quantidade de processos suspensos no STJ.



Em um cenário composto por todas as instâncias do poder judiciário, dos 286.834 processos em tramitação no STJ, as 27 afetações, de temas repetitivos, ocorridas entre julho de 2020 até agosto de 2021, impactaram em 182.610 processos sobrestados. No mesmo período, foram julgados 28 temas repetitivos, ocasionando um impacto em 24.226 processos julgados<sup>43</sup>.

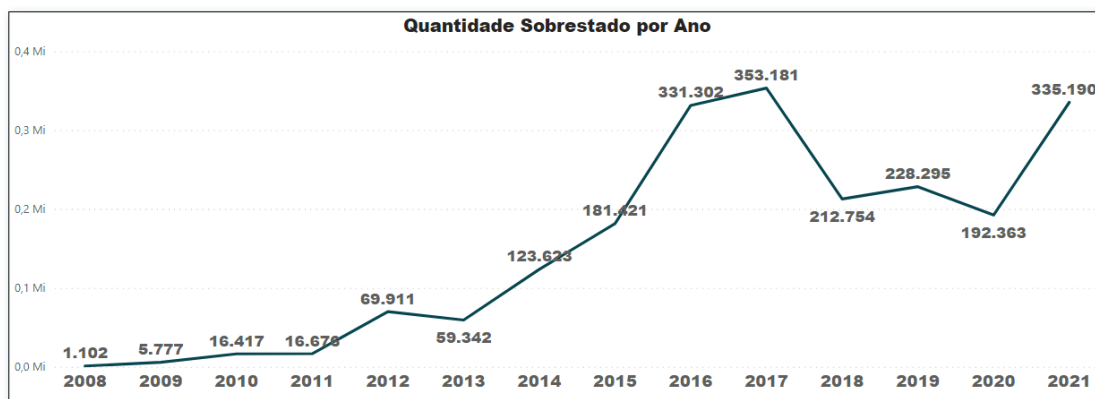
Os Gráficos 12 e 13 ilustram o histórico de temas cadastrados na base de dados de precedentes qualificados do STJ, nos últimos 13 anos e o número de processos impactados, em todas instancias do judiciário, neste mesmo período:

Gráfico 12- Série histórica da quantidade de temas cadastrados na base de dados de precedentes qualificados do STJ.



<sup>43</sup> CONSELHO NACIONAL DE JUSTIÇA. Disponível em: <https://www.cnj.jus.br/pesquisas-judiciarias/demandas-repetitivas/>. Acesso em: 01 jan. 2022.

Gráfico 13- Número de processos sobrestados em todas instancias do judiciário.



Os gráficos apontam uma redução na quantidade de temas cadastrados nos últimos 12 anos, no Superior Tribunal de Justiça. Inversamente, percebe-se o crescimento no número de processos sobrestados, aguardando a decisão daquelas teses. Este cenário deve-se principalmente à complexidade intrínseca do rito dos repetitivos no STJ. Inserida a partir do ano de 2016, as regras impostas pelo Código de Processo Civil (CPC) e Regimento Interno (RI) daquela Corte, juntamente aos estudos realizados na criação de novos Temas repetitivos, impactaram diretamente na quantidade de Temas cadastrados, o que também refletiu no caráter qualitativo do rito, gerando uma relação inversa entre a quantidade de temas cadastrados e o número de processos sobrestados.

## 4.2 O sistema Athos no apoio a criação de novos temas repetitivos

Em setembro de 2016 entrou em funcionamento o Núcleo de Gerenciamento de Precedentes (NUGEP), considerado, ferramenta fundamental para a consolidação do sistema de precedentes judiciais trazido pelo novo Código de Processo Civil. A criação da unidade foi mais um passo do STJ para se adequar às inovações legislativas que prometem, com o reforço do precedente jurisprudencial, aumentar a segurança jurídica e reduzir a morosidade dos processos.

A criação deste núcleo partiu da determinação da Resolução 235/16, do Conselho Nacional de Justiça (CNJ), núcleos semelhantes foram instalados em todos os tribunais superiores, nos tribunais estaduais, nos regionais federais e do trabalho.

Atualmente, conhecido como NUGEPNAC, o Núcleo de Gerenciamento de Precedentes é uma unidade técnico-administrativa encarregada de controlar informações sobre os processos que geram precedentes, como os recursos repetitivos e os incidentes de assunção de competência. Entre os seus objetivos está o melhor gerenciamento dos processos e das respectivas informações, e o aperfeiçoamento da maneira de julgar.

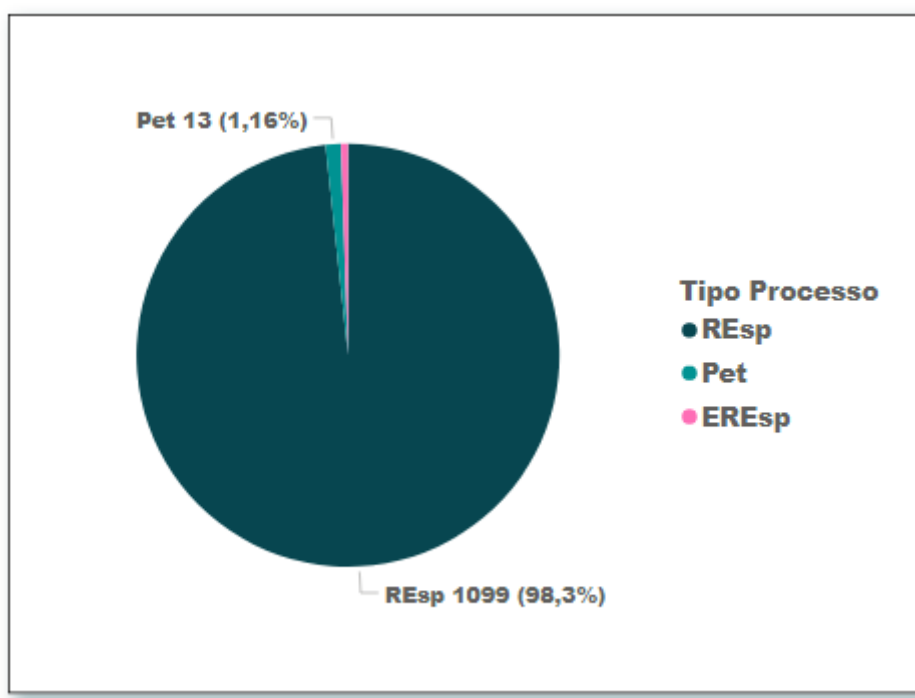
O NUGEPNAC tem a responsabilidade de identificar teses repetitivas, em conjunto com a Comissão Gestora de Precedentes e de Ações Coletivas (COGEPAC), a qual é formada por três ministros do STJ, como membros efetivos e um ministro como suplente e um juiz supervisor, representantes das seções de direito público, direito privado e direito penal.

O sistema Athos é utilizado pela COGEPAC, desde setembro de 2019, no monitoramento das teses repetitivas no STJ. A partir dos grupos de processo sugeridos pelo Athos, a equipe técnica do Núcleo inicia os estudos para identificar se aquela tese mapeada pelo sistema de IA será candidata a um novo tema repetitivo no âmbito do Superior Tribunal de Justiça, ao identificar, antes mesmo da distribuição, os processos com multiplicidade de uma mesma questão de direito, o sistema amplia o número de recursos candidatos à afetação para julgamento na sistemática dos repetitivos. A ferramenta também monitora e aponta processos com entendimentos convergentes ou divergentes entre os órgãos fracionários do STJ, processos com matéria de notória relevância e, ainda, possíveis distinções ou superações de precedentes qualificados. (NORONHA, 2020).

No presente estudo, usaremos apenas dados obtidos nos anos de 2020 e 2021, pois entendemos que os dados do ano de 2019 não são suficientes para fins comparativos, pois foram lançados somente a partir do mês de setembro. Ademais, antes desse marco, o NUGEPNAC não mantinha uma rotina de registro de indicadores.

O Sistema Athos possui diversas possibilidades de critério de agrupamento de processos, dentre eles: por ministro, por deslocamento, por órgão julgador ou por classe de processo. Nesse sentido, ressalta-se que internamente os estudos do Núcleo são direcionados aos Recursos Especiais (REsp) e aos Agravos em Recurso Especial (AREsp). Isto pode ser constatado na base de dados de precedentes qualificados, onde, dos 1124 temas cadastrados no STJ, 98,3% (1099) dos processos paradigmas afetados são REsp's, e 1,18% são os demais (Pet e EREsp), como observado no Gráfico 14<sup>44</sup>:

Gráfico 14- Tipos de RRC cadastrados na base de dados do STJ.



Um dos métodos de agrupamento mais utilizado pelos usuários do NUGEPNAC é o agrupamento antes da distribuição dos processos, em consonância com o Art. 2º, I, da Portaria STJ/GP n. 299 de 19 de julho de 2017:

[...] Art. 2º Ficam delegadas ao presidente da comissão as seguintes competências:

I – Despachar, antes distribuição, em recursos indicados pelos Tribunais de origem como representativo da controvérsia; [...]

<sup>44</sup> SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: <  
[https://processo.stj.jus.br/repetitivos/temas\\_repetitivos/](https://processo.stj.jus.br/repetitivos/temas_repetitivos/)>. Acesso em: 10 set. 2021.



Neste módulo, o usuário indica o critério de pesquisa e o Athos agrupa os processos de acordo com as teses semelhantes encontradas. Assim, o operador do sistema, em sua tela de customização, configura o *output* para resultados agrupados com no mínimo 50 processos. A similaridade exigida pelo Núcleo é uma acurácia de mínimo 90%, e que o sistema apresente ao menos um processo como paradigma (Recurso Representativo da Controvérsia-RRC).

O RRC é o processo escolhido dentre vários outros que possuam a mesma questão de direito, e que servirá como caso concreto paradigma para que o Superior Tribunal de Justiça fixe a tese jurídica, tornando-a tema repetitivo.

Por esta razão, a lei e a prática forense exigem que sua escolha seja criteriosa e pautada nos requisitos da existência de multiplicidade de recursos com idêntica questão de direito e a escolha de mais de um processo, com a maior gama de fundamentos e argumentos, para tramitação conjunta no STJ.

Os recursos representativos de controvérsia poderão advir de quatro situações<sup>45</sup>:

- Recursos admitidos pelos tribunais de origem com fundamento no art. 1.036, § 1º do CPC de 2015. Ao admitir RRCs, o tribunal de origem inicia o controle para sobrestamento de demais processos no estado ou na região em que discutida a mesma matéria.
- Recursos admitidos pelos tribunais de origem com fundamento no art. 1.041 do CPC que, tratam de matéria que já foi firmada como tese repetitiva por tribunal superior, e não houve sua aplicação pelo tribunal de origem, em juízo de retratação, considerando-se possível distinção ou superação de tese. Ao admitir RRCs, o tribunal de origem

---

<sup>45</sup>SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: <<https://www.stj.jus.br/sites/portalp/Precedentes/informacoes-gerais/controversias>>. Acesso em: 23 de nov. 2021

inicia o controle para sobrestamento de demais processos no estado ou na região em que discutida a mesma matéria.

- Recursos encaminhados ao STJ que não foram selecionadas como representativas da controvérsia pelos tribunais de origem, mas que, por meio do desenvolvimento de trabalho de inteligência no STJ, antes da distribuição, foram identificados por conterem matérias com “potencial de repetitividade ou com relevante questão de direito, de grande repercussão social, aptas a serem submetidas ao Superior Tribunal de Justiça, sob a sistemática dos recursos repetitivos” (art. 46-A, IV, do RISTJ).
- Recursos especiais interpostos contra o julgamento de mérito de um Incidente de Resolução de Demandas Repetitivas (IRDR) pelos tribunais de segunda instância (art. 256-H, do RISTJ). Isso porque a tese jurídica firmada pelo STJ, ao apreciar o mérito desse recurso, terá aplicação nacional mesmo que não seja objeto de afetação.

Como dito, a similaridade exigida na busca por novos processos pelo NUGEPNAC é de no mínimo 90%, e de acordo com a Assessoria de Inteligência Artificial (AIA), esta busca parte da posse de um processo e uma peça processual (número de registro e a página). Assim o operador no núcleo indica na tela de pesquisa à similaridade desejada e o processo paradigma de pesquisa. Se o operador do sistema assim desejar, poderá buscar por um processo novo, ou por um processo que já possua decisão, simulando uma pesquisa de jurisprudência.

Figura 3- Tela do filtro de busca por processos semelhantes do sistema Athos.

**Filtros de Busca por Processos Semelhantes**

A pesquisa se limita aos processos tramitando e os que receberam decisão após 01/01/2016.

Buscar Semelhantes Sem Decisão  Buscar Antecedentes

Nº Registro ou Classe e Número: \*

Primeira folha (e-STJ) da Peça:

Similaridade %:

Ministro Relator:

Órgão Julgador:

Decisão:

Local (último deslocamento do processo):

Classe:

Tramitando:

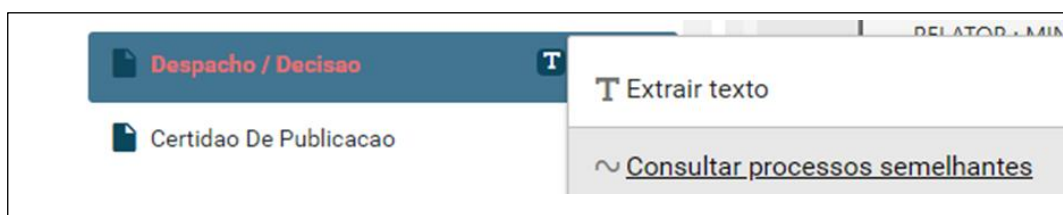
Critérios de Pesquisa Textual:

Limpar

Fonte: Assessoria de Inteligência Artificial do STJ.

O sistema Athos também permite selecionar no menu do visualizador de processos a opção “Consultar processos semelhantes”.

Figura 4- Tela consultar processos semelhantes do sistema Athos.



Fonte: Assessoria de Inteligência Artificial do STJ.

Independente do caminho usado para chegar na pesquisa, o operador do sistema pode complementar alguns parâmetros de filtros como “Ministro Relator”, “Deslocamento”, “Órgão Julgador”, dentre outros, para que sejam retornados apenas processos que correspondam a esse filtro.

O sistema identifica pela página do processo qual o vetor que deverá ser comparado, quais vetores estão próximos dele e a quais processos

pertencem a esses vetores, bem como se os dados do processo correspondem aos filtros aplicados.

O resultado será apresentado com a indicação da similaridade semântica entre a peça do processo indicado na pesquisa e a peça do processo encontrado, O sistema também informa outros dados como: a página da peça encontrada, status do processo, deslocamento e se existem decisões ou acórdãos publicados para os processos encontrados.

O resultado retornado pelo sistema é ordenado a partir dos processos de maior similaridade aos de menor similaridade. A figura 5 demonstra a tela do sistema Athos com o resultado de uma busca por similaridade.

Figura 5- Tela do sistema Athos com o resultado busca por similaridade.

**STJ** Sistema Justiça - Athos

Total de Processos: 88  
100 peças encontradas.

**Processos**

REP. 1210170 / TR...  
2019/0141716-5

Relator(a): MIN. CÂNDIDO MENEZES

Órgão Julgador: PRIMEIRA TURMA

Origem: TRIBUNAL REGIONAL FEDERAL DA 1ª REGIÃO

Assunto: Apreensão

Status: Processo baixado

Resumo Indicativo: AMBIENTAL MANDADO DE SEGURANÇA DESTRUIÇÃO DE VEGETAÇÃO NATIVA SEM AUTORIZAÇÃO DO ÓRGÃO COMPETENTE APREENSÃO DE VEÍCULO NOMEAÇÃO DO PROPRIETÁRIO DO BEM COMO FIEL DEPOSITÁRIO DECRETO 6514/2008 POSSIBILIDADE SENTENÇA PARCIALMENTE REFORMADA

Similaridade: 97,88

Primeira Folha (e-STJ): 340

Nº Registro Paradigma/Folha: 2019/0000000 / 220

Saída para STJ - Processo eletrônico baixado e recebido em 03/03/2019 10:42:02

Decisão Desfecho: Proferido despacho de mero expediente determinando a devolução dos autos

REP. 1011011 / TR...  
2019/0141716-5

Relator(a): MIN. MARCO ANTONIO DE MELLO

Relator(a) Anterior: MIN. PRESIDENTE DA COMISSÃO GESTORA DE PRECEDENTES

Órgão Julgador: SEGUNDA TURMA

Origem: TRIBUNAL REGIONAL FEDERAL DA 5ª REGIÃO

Assunto: Ambiental

Status: Processo baixado

Resumo Indicativo: ADMINISTRATIVO INFRAÇÃO AMBIENTAL TRANSPORTE MADEIRA APREENSÃO DE VEÍCULO IMPOSSIBILIDADE

Similaridade: 97,74

Primeira Folha (e-STJ): 427

Nº Registro Paradigma/Folha: 2019/0000000 / 356

Saída para STJ - Processo eletrônico baixado e recebido em 03/03/2019 10:42:02

Acórdão Desfecho: Deferindo

Fonte: Assessoria de Inteligência Artificial do STJ.

A pesquisa textual é outro recurso utilizado na busca por peças processuais disponíveis, com o auxílio de operadores de pesquisa (ADJ, PROX, COM, E, OU, NÃO).

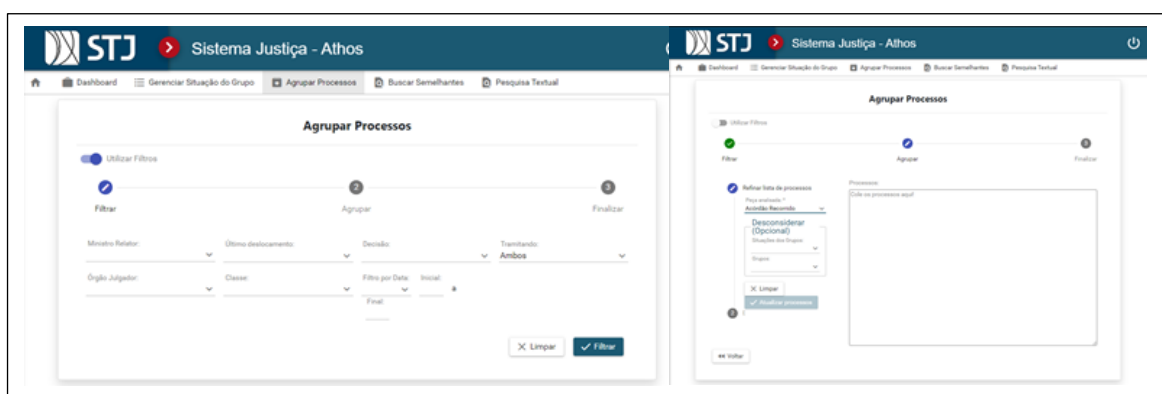
Figura 6- Tela da busca textual do sistema Athos.



Fonte: Assessoria de Inteligência Artificial do STJ.

Outra forma de pesquisa no Athos é a possibilidade de busca por semelhança, com o RRC já definido. Neste modelo, alimenta-se o sistema com um processo direcionador do Recurso Representativo da Controvérsia e o Athos cria grupos de processos, de acordo com aquela tese apontada no *input* do sistema. Dessa forma, o usuário escolhe um conjunto de processos, indicando o número do registro ou um conjunto de filtros como: processos deslocados para um determinado gabinete, sob a relatoria de um determinado Ministro, que estão ou não tramitando, dentre outros parâmetros possíveis, como vemos abaixo:

Figura 7- Tela busca por semelhança do sistema Athos.



Fonte: Assessoria de Inteligência Artificial do STJ.

Neste módulo, o operador indica o tipo da peça processual que será analisada em cada processo, por exemplo “Acórdão Recorrido”, bem como a similaridade mínima desejada. A comparação é feita entre as peças de cada processo com todas as outras peças do conjunto de processos informados. O

agrupamento de processos, ou peças similares, ocorre ao identificar em uma lista de vetores quais estão próximos dentro de uma similaridade indicada.

O sistema buscará o endereço vetorial das peças de cada processo informado, ao comparar as distâncias vetoriais entre elas, ou seja, identifica dentre as peças selecionadas quais são vizinhas entre si no universo vetorial do Athos, com uma distância que represente a similaridade indicada. Posteriormente, durante os estudos, a equipe técnica do NUGEPNAC agrupa outros processos e verifica a possível multiplicidade de processos, através da função “Monitoramento”. Esta função permite que após um agrupamento realizado, pode-se retirar do grupo alguns processos sugeridos, incluir novos processos ou até indicar processos que representam o que não se deseja para esse grupo, ou seja, exemplos negativos para o grupo.

Os analistas da AIA esclarecem que ao se indicar o monitoramento de um grupo, outros parâmetros podem ser definidos como: indexação legislativa, classes processuais e critérios textuais como a existência ou não de determinados termos.

Figura 8- Tela busca monitoramento de grupos do sistema Athos.

The screenshot displays the 'Editar Grupo' interface for 'AIA - Descarte Acórdão'. The header includes the STJ logo and navigation options like 'Dashboard', 'Gerenciar Situação do Grupo', 'Agrupar Processos', 'Buscar Semelhantes', and 'Pesquisa Textual'. The main form contains the following fields and options:

- Código:** AIA
- Nome:** AIA
- Peça Analisada:** Todas
- Local:** SECRETARIA DE INFORMÁTICA
- Similaridade %:** 95
- Situação:** Concluído
- Ramo de Direito:** (dropdown menu)

**Adicionais**

Informações adicionais:  
AIA

Critério de Pesquisa de Jurisprudência:

Quantidade RRC: 0

Classe Processual: (dropdown menu)

RRC

Monitoramento Ativo

Visível a outras Unidades

Visível no escaninho

Curadoria Externa

**Critérios Adicionais de Monitoramento (Opcionais)**

Ao selecionar algum dos critérios opcionais, o sistema refinará o monitoramento do grupo, exigindo que o processo, além de ter peça com similaridade mínima à escolhida para o grupo, preencha os critérios selecionados. Atenção: Caso algum dos requisitos não seja cumprido, o processo não será inserido no grupo. Ex: Além de ter uma peça de Acórdão Recorrido com mais de 90% semelhante, o processo obrigatoriamente deverá ter como origem o Tribunal de Justiça do Estado de Alagoas e deverá ser da parte Município de Maceió

Contém palavra ou expressão:  i

Texto:

Não Contém palavra ou expressão:  i

Texto:

Contém Indexação Legislativa:  i

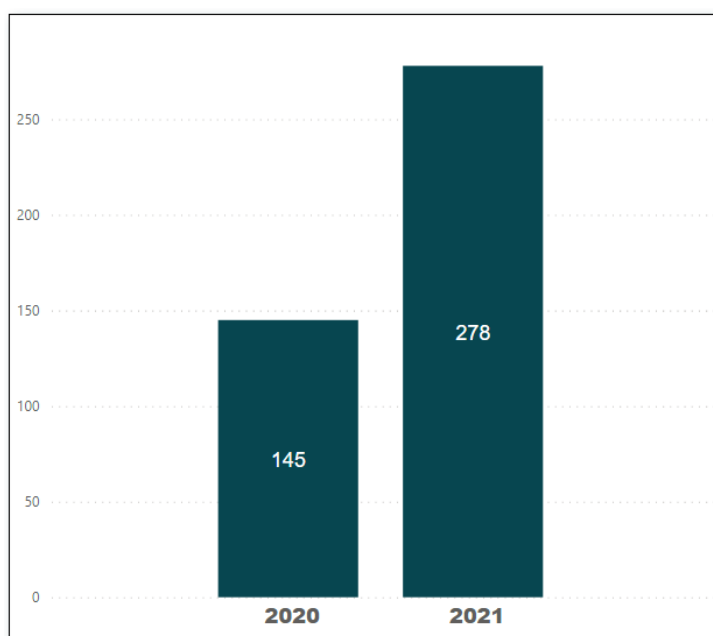
Indexação Legislativa:

Fonte: Assessoria de Inteligência Artificial do STJ.

A ferramenta possui também a funcionalidade “Rodar Acervo”, onde o usuário cria um grupo, manualmente ou pelo agrupamento, sendo possível atualizar o grupo com processos semelhantes aos do próprio grupo e que estejam no seu local de trabalho, sem realizar um novo agrupamento. Dessa forma, o usuário pode usar a funcionalidade, a qual consiste em identificar no acervo do ambiente de trabalho local os processos semelhantes aos processos marcados como paradigmas no grupo que está sendo analisado.

Uma outra forma de agrupamento disponibilizado pelo Athos é a busca por paradigmas já consolidados na jurisprudência do STJ. O operador alimenta o sistema com 10 (dez) paradigmas sobre uma tese já pacificada e o sistema cria grupos daquele modelo, com vistas a padronização e uma futura possibilidade de julgamento daquela tese no rito dos repetitivos, e conseqüentemente, mitigar excessivos julgamentos monocráticos sobre a demanda destacada. Atualmente, o Núcleo de Gerenciamento de Precedentes do STJ possui 494 grupos em monitoramento, sendo que houve um aumento de 91,74% na criação de grupos de estudos, saltando de 145 (2020) para 278 (2021), como demonstrado no Gráfico 15:

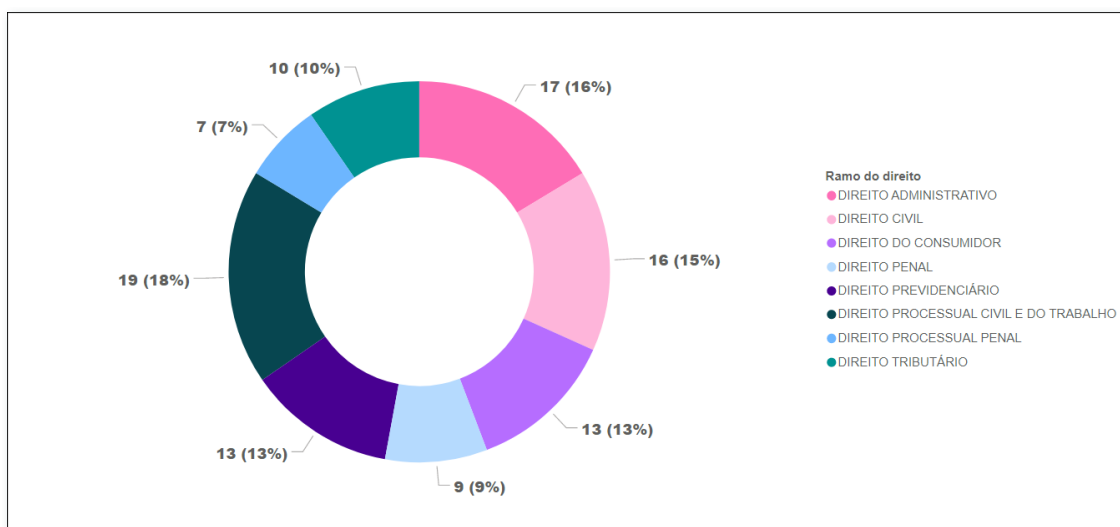
Gráfico 15- Quantidade de grupos monitorados pelo NUGEPNAC.



Após a criação dos grupos com auxílio do Athos, por meio das metodologias citadas, os operadores do NUGEPNAC iniciam os estudos dos

grupos. Partem pela escolha das peças processuais que serão estudadas, que em maioria são REsp's e Acórdãos. Em seguida, definem o ramo do direito em que se enquadra o grupo analisado. De acordo com os dados de precedentes qualificados, no período de 2019 a 2021, o Direito Processual Civil e do Trabalho é o ramo do direito com maior índice de processos estudados pelo Núcleo, com 18%. Em seguida aparecem o Direito Administrativo 16%, Direito Civil 15%, o Direito Previdenciário e Direito do Consumidor com 13%, e os demais ramos do Direito como demonstrado abaixo<sup>46</sup>:

Gráfico 16- Ramos do Direito nos processos estudados pelo NUGEPNAC.



Não obstante a relevância quantitativa proporcionada pelo rito dos Recursos Repetitivos no ordenamento jurídico brasileiro, descrita neste capítulo do trabalho, a COGEPAC entende que a qualidade dos temas definidos por aquela Corte é de grande relevância. Nesse sentido, buscam priorizar, em seus estudos, os princípios da celeridade na tramitação dos processos, a isonomia de tratamento às partes processuais e a segurança jurídica.

Para alcançar este objetivo, os operadores do NUGEPNAC, analisam todos os RRC's remetidos ao STJ pelos tribunais de origem, pelas entidades conveniadas, ou selecionados no próprio STJ, quanto a aspectos formais, dentre eles, os requisitos de admissibilidade do Recurso Especial. Aos quais podemos destacar:

<sup>46</sup> SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: <[https://processo.stj.jus.br/repetitivos/temas\\_repetitivos/](https://processo.stj.jus.br/repetitivos/temas_repetitivos/)>. Acesso em: 20 jan. 2022.



- I. Requisitos de admissibilidade comuns a outros tipos de recursos- interesse da parte (Art. 17, CPC), a legitimidade do pedido (Art. 17, CPC), a inexistência de impedimento e suspeição (Art. 144 e 145, CPC), o preparo (Art. 1.007, CPC) e a tempestividade (Art. 1,030, CPC);
- II. Requisitos previstos na Constituição Federal, no seu Art. 105, III- julgar, em recurso especial, as causas decididas, em única ou última instância, pelos Tribunais Regionais Federais ou pelos tribunais dos Estados, do Distrito Federal e Territórios, quando a decisão recorrida: a) contrariar tratado ou lei federal, ou negar-lhes vigência; b) julgar válido ato de governo local contestado em face de lei federal; c) der a lei federal interpretação divergente da que lhe haja atribuído outro tribunal.
- III. Existência de prequestionamento- Sumulas 211/STJ: “Inadmissível recurso especial quanto à questão que, a despeito da oposição de embargos declaratórios, não foi apreciada pelo Tribunal a quo” e 282/STF: “É inadmissível o recurso extraordinário, quando não ventilada, na decisão recorrida, a questão federal suscitada”.
- IV. Não discutir fatos e direito, súmula 7/STJ- “A pretensão de simples reexame de prova não enseja recurso especial”.
- V. Esgotamento dos recursos ordinários, súmula 207/STJ- “É inadmissível Recurso Especial quando cabíveis embargos infringentes contra o acórdão proferido no tribunal de origem”.
- VI. Fundamentação deficiente, súmula 284/STF- “É inadmissível o recurso extraordinário, quando a deficiência na sua fundamentação não permitir a exata compreensão da controvérsia”.
- VII. Art. 1.036, § 6º do CPC- “Somente podem ser selecionados recursos admissíveis que contenham abrangente

argumentação e discussão a respeito da questão a ser decidida”.

Assim, com vistas a garantir RRC's aptos, a COGEPAC inclui, nos estudos dos grupos criados pelo sistema Athos, a análise da multiplicidade de recursos, prevista no Art. 1.036 do CPC:

[...] Art. 1.036 Sempre que houver multiplicidade de recursos extraordinários ou especiais com fundamento em idêntica questão de direito, haverá afetação para julgamento de acordo com as disposições desta Subseção, observado o disposto no Regimento Interno do Supremo Tribunal Federal e no do Superior Tribunal de Justiça [...].

Este é um destaque do sistema Athos, pois antes de sua implantação, no ano de 2019, o mapeamento e a busca por recursos idênticos no STJ eram feitos de forma artesanal, por amostragem, com auxílio de planilhas eletrônicas. O núcleo responsável, há época, não possuía recursos tecnológicos que pudessem dar suporte ao complexo desafio de localizar, dentre os milhares de processos que ingressam anualmente naquela Corte, um tema que se repetisse frequentemente, e a busca, logo na entrada processual de outros recursos semelhantes, a fim de satisfazer o Art. 1.036 do CPC em consonância com a multiplicidade de recursos.

A COGEPAC inclui ainda nos seus estudos critérios como:

- relevância da matéria;
- quantidade de julgados nos órgãos fracionários (Turmas, Seções e Corte Especial);
- repercussão social da tese;
- existência de pacificação da tese jurídica em análise;
- sedimentação da jurisprudência como pressuposto para afetação do recurso especial ao rito dos repetitivos;
- uniformização e consolidação da jurisprudência.

Após os estudos realizados pela equipe do NUGEPNAC, o Presidente da Comissão Gestora de Precedentes e de Ações Coletivas, nos

termos dos art. 256-B a 256-D e da Portaria STJ/GP n. 98/2021, por meio de despacho nos autos<sup>47</sup>:

- a. destaca a questão de direito delimitada pelo tribunal de origem indicada como repetitiva no estado ou na região, com o registro de informações importantes do processo;
- b. determina a abertura de vista ao Ministério Público Federal para manifestação a respeito dos pressupostos de admissibilidade do recurso especial como representativo da controvérsia;
- c. determina a distribuição do recurso a um dos ministros do STJ.

Após a conclusão dos autos ao relator sorteado, o ministro irá estabelecer a controvérsia da tese em questão, e posteriormente, decidir no prazo estabelecido no art. 256-E, do Regimento Interno do STJ, de 60 dias úteis, sob pena de rejeição tácita, se propõe a afetação do recurso representativo da controvérsia ao rito dos repetitivos.

Santos, Cunha e Filho (2020) esclarecem ainda que pode o relator, no tribunal superior, selecionar outros recursos sobre a controvérsia repetitiva, para servirem de “modelo” ou “piloto”. Pois, pode haver outros casos com contraditório mais aprofundado, ou com mais argumentos que se somem, ou com mais claro manancial fático-jurídico, os quais devem ser somados, para enfrentamento na motivação do julgamento final. Esse julgamento final, deverá fornecer resposta jurisdicional a toda a controvérsia repetitiva. Os autores enfatizam que de acordo com o CPC/2015, tratar de toda a controvérsia, quer dizer, não apenas a posta em um único recurso – art. 1.036, §§ 1º e 4º – de modo que se trata, etiologicamente, de um grande contraditório nacional que enfeixe todos os argumentos existentes sobre o assunto, livre da angustura talvez sobrevivente do fascínio jurisdicional de possível prequestionamento –

---

<sup>47</sup> SUPERIOR TRIBUNAL DE JUSTIÇA: Disponível em: <https://www.stj.jus.br/sites/portalp/Precedentes/informacoes-gerais/controversias>. Acesso em: 02 dez. 2021.

incompatível com o sistema de julgamento de macro-lides em recursos repetitivos.

Nesta fase do rito de criação de tema repetitivo, o Athos torna-se fundamental, pois possui capacidade de agrupar novos processos de acordo com a controvérsia apontada pelo relator. Sua acurácia de no mínimo 90%, possibilita que a COGEPAC busque processos com alto grau de semelhança daquela controvérsia já definida, reforçando o intuito da valoração qualitativa que o sistema Athos proporciona à todas as fases dos estudos de potenciais novos temas de recursos qualificados.

A Tabela 3 demonstra a evolução das controvérsias criadas no NUGEPNAC, durante os anos de 2020 e 2021, nela encontra-se os dados comparados entre controvérsias criadas com o auxílio do Athos e as chamadas controvérsias globais. Essas, são controvérsias criadas a partir dos RRC enviados pelos tribunais de origem, sugeridas por outros órgãos oriundos de acordos de cooperação técnica, ou sugeridas pelos próprios ministros do STJ, ou seja, controvérsias criadas sem o auxílio da IA. Percebe-se que o sistema Athos foi responsável pela criação de 153 controvérsias e as controvérsias globais totalizaram 229 no período de 2020 e 2021.

Tabela 3- Quantidade de controvérsias criadas entre os anos de 2020 e 2021 pelo NUGEPNAC.

Ano	Mês	Controvérsias Globais	Controvérsias do Athos
2020	janeiro	0	0
2020	fevereiro	9	6
2020	março	11	6
2020	abril	9	7
2020	maio	10	8
2020	junho	7	4
2020	julho	0	0
2020	agosto	14	10
2020	setembro	4	2
2020	outubro	3	1
2020	novembro	7	6
2020	dezembro	27	18
2021	janeiro	0	0
2021	fevereiro	1	1
2021	março	12	8
2021	abril	12	10
2021	maio	8	6
2021	junho	14	12
2021	julho	4	2
2021	agosto	12	8
2021	setembro	9	8
2021	outubro	18	14
2021	novembro	21	8
2021	dezembro	17	8
<b>Total</b>		<b>229</b>	<b>153</b>

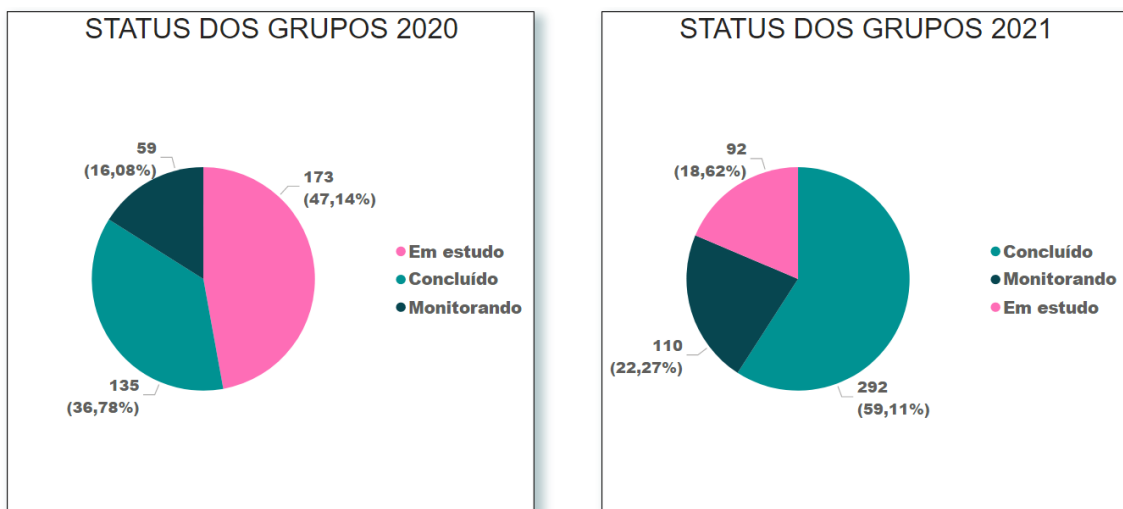
Ressalta-se que no período mencionado, a quantidade de controvérsias criadas pelo sistema de IA do STJ passaram de 68 em 2020, para 85 em 2021. Isto é, um aumento de 25% no período de um ano. Destaca-se também que o Athos já é responsável pela criação de 40,05% controvérsias dos estudos do COGEPAC.

O NUGEPNAC possui uma rotina de monitoramento de todos os grupos criados pelo sistema Athos. Até dezembro 2021, o núcleo possuía 494 grupos em observação, os quais estão divididos em três *status*: “Monitorando”, “Em Estudo” e “Concluído”, sendo que:

- Em Estudo- são grupos que não possuem Recurso Representativo da Controvérsia definido;
- Monitorando- grupos de processos que já possuem RRC definido. Porém, precisam passar pelo crivo do Ministério (MP) ou ministros da Comissão Gestora de Precedentes, possuem RRC, mas estão em uma fase pré-distribuição;
- Concluído- já foram ao MP, já foram distribuídos, já viraram controvérsia e tema. Não necessita de monitoramento.

Atualmente, existem 110 grupos no status “monitorando”, 292 “concluído” e 92 “em estudo” no NUGEPNAC. No entanto, em 2020 havia 59 grupos no status “monitorando”, 135 “concluído” e 173 “em estudo”.

Gráfico 17- Status dos grupos monitorados nos anos de 2020 e 2021 pelo NUGEPNAC.



Percebe-se uma redução de 46,82% dos grupos “em estudo” e o aumento de 86,44% dos grupos “monitorando”. Isso se deve, em parte, ao complexo andamento do rito dos repetitivos. Após a identificação do RRC, o estabelecimento da tese e da controvérsia, por parte da Comissão Gestora de Precedentes e de Ações Coletivas, o trâmite previsto do Art. 256 ao Art. 256-M do Regimento Interno do STJ possui diversos procedimentos, com decisões subjetivas, que de certa forma, impedem ou inviabilizam que aquela tese monitorada transforme-se em um tema em Recurso Repetitivo (RR). A própria ministra do STJ: Nancy Andrighi, já evidenciou a relevância de se estudar o tema dos incidentes de resolução de demandas repetitivas (IRDR) por meio de seminários e audiências públicas, e enfatizou a dificuldade de se definir a natureza jurídica do IRDR, em virtude da falta de clareza legal<sup>48</sup>.

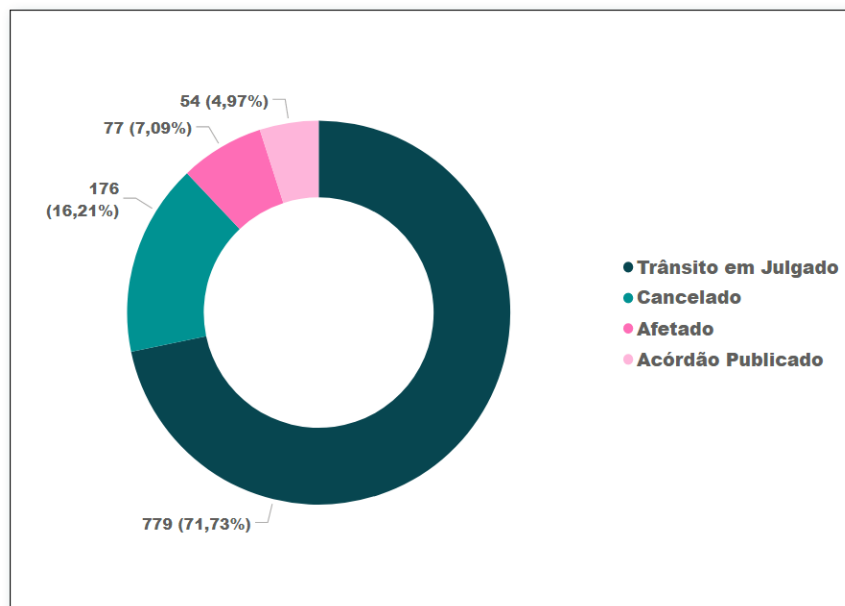
Uma evidência do quão complexo é a conclusão do processo de criação de temas repetitivos, está demonstrada na quantidade de temas cancelados na base de dados do STJ. Onde, após percorrerem todas as etapas de estudos, descritas anteriormente, e já publicado o(s) acórdãos de afetação, 176 temas cadastrados foram cancelados, ou seja, 16,21% dos temas cadastrados. Como verificado abaixo<sup>49</sup>:

---

<sup>48</sup> SUPERIOR TRIBUNAL DE JUSTIÇA: Disponível em: <[https://intranet.stj.jus.br/SGL/jsps/noticias\\_internas/noticia\\_detalhe.jsp?seq\\_noticia=33167](https://intranet.stj.jus.br/SGL/jsps/noticias_internas/noticia_detalhe.jsp?seq_noticia=33167)>. Acesso em: 02 dez 2021

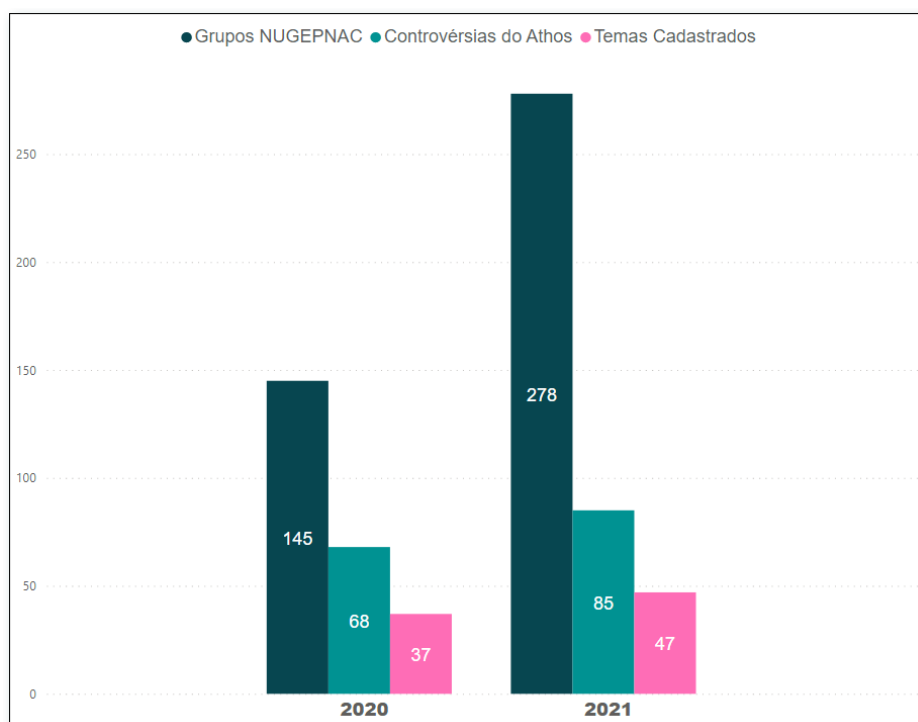
<sup>49</sup> SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: <<https://processo.stj.jus.br/repetitivos/temasrepetitivos/>>. Acesso em: 02 dez 2021

Gráfico 18- Quantidade de Temas cancelados cadastrados na base de dados do STJ.



Mesmo diante da complexidade apontada, os dados disponibilizados pela equipe do NUGEPNAC evidenciam o aumento no número de novos temas criados entre os anos de 2020 e 2021. O sistema Athos obteve aumento de 91,72% no número de grupos criados, enquanto a criação de controvérsias subiu 25%, e os temas criados tiveram um acréscimo de 27,02% no mesmo período.

Gráfico 19- Quantidade de Grupos/Controvérsias/Temas criados pelo sistema Athos entre os anos de 2020 e 2021



Dados disponibilizados pela equipe da Assessoria de Inteligência Artificial do STJ confirmam esta afirmativa, pois apontam que a produtividade do sistema Athos nos módulos “Busca por Similaridade”, “Busca Textual”, “Agrupamento”, “Monitoramento” e “Rodar Acervo” mantiveram-se em alta no período descrito anteriormente. Esses dados encontram-se armazenados no banco de dados da aplicação e complementam as informações disponibilizadas pela equipe do NUGEPNAC, pois, são registros de requisições ao sistema Athos que não foram tratadas como oficiais pela equipe do Núcleo, ou foram apenas usadas no apoio aos seus estudos e pesquisas ou simplesmente descartadas. Porém, por constarem no banco de dados como requisições realizadas no sistema, optamos em apresentá-las, com objetivo de relatar a atual eficiência da ferramenta ao disponibilizar dados comparativos para novos estudos.

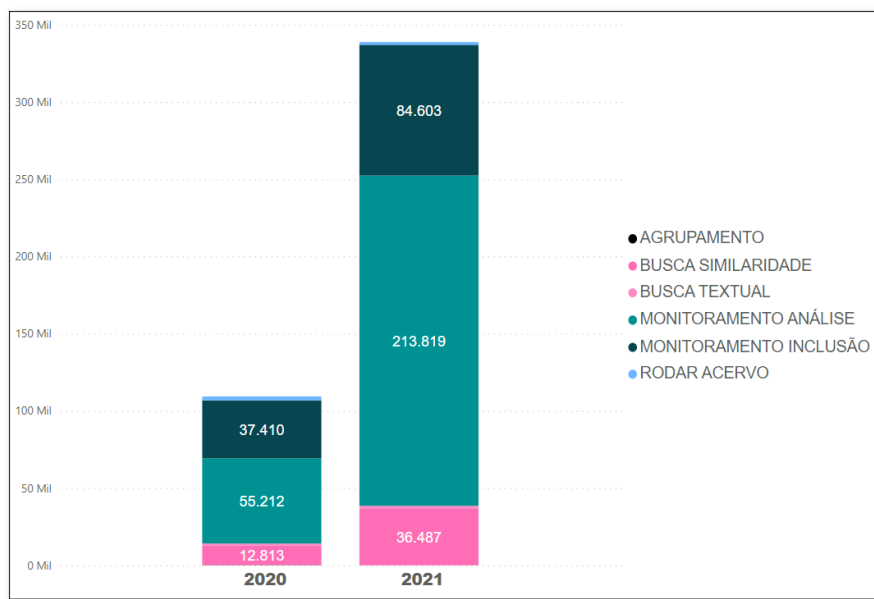
Para tanto, apresentaremos os dados das principais funcionalidades utilizadas pelos operadores do sistema de IA partindo da funcionalidade “Monitoramento”, a qual pelas regras da arquitetura do sistema definida pelos desenvolvedores da equipe da assessoria de inteligência artificial subdivide-se em:

- Monitoramento Análise: onde todos novos processos são comparados com grupos monitorados em gabinetes, Jurisprudência, Secretaria Judiciária etc.
- Monitoramento Inclusão: processos que foram inseridos em grupos monitorados por atenderem aos critérios de similaridade e filtros de metadados.

O gráfico 20 ilustra a quantidade de requisições das funcionalidades “Monitoramento” do Athos, nos anos de 2020 e 2021. Somente a funcionalidade “monitoramento inclusão” obteve um aumento de 126,15% na quantidade de execuções realizadas, passando de 37.410, em 2020, para 84.603 em 2021. A outra funcionalidade “Monitoramento Análise”, também apresentou resultados crescentes, ao partir de 55.212 monitoramentos realizados, para 213.819, ou seja, um acréscimo de 287,27%.



Gráfico 20- Quantidade de funcionalidades utilizadas no sistema Athos entre os anos de 2020 e 2021.



O gráfico também demonstra os dados dos recursos de busca. Nesse sentido, o Athos possui duas opções: “Busca Similaridade” e “Busca Textual”. A “Busca Textual” consiste na pesquisa com uso de critérios textuais simples, enquanto a “Busca Similaridade” permite a pesquisa a partir de um processo paradigma. Assim, os dados disponibilizados pela AIA, apontam que foram efetuadas 52.489 buscas pelos operadores do Athos, entre os anos de 2020 e 2021, sendo 49.300 por similaridade e 3.189 textual. O Gráfico 17 aponta um aumento de 184,76% nas buscas por similaridade no período, enquanto as buscas textuais saíram de 1.158 em 2020 para 2.031 em 2021, isto é, aumento de 75,38%.

Por fim, duas funcionalidades obtiveram redução nos seus números de execuções monitoradas pela AIA. A função “Rodar Acervo” foi utilizada 4.406 vezes, com redução de 30,21% no período mencionado. Já a funcionalidade “Agrupamento” foi utilizada 254 vezes, com redução 52,35% entre os anos de 2020 e 2021, como visto no Gráfico 21:

Gráfico 21- Funcionalidades Rodar Acervo/Agrupamentos utilizadas no sistema Athos entre os anos de 2020 e 2021.



Em seu protagonismo, no apoio à seleção de recursos aptos, candidatos à temas em recurso repetitivo, o Athos demonstrou crescente evolução, pois os *logs*<sup>50</sup>, registrados em seu banco de dados, disponibilizados pela AIA do STJ, demonstram seu desempenho na criação de grupos e identificação de processos paradigmas. A quantidade de processos envolvidos nessas operações também evoluiu consideravelmente entre os anos de 2020 e 2021.

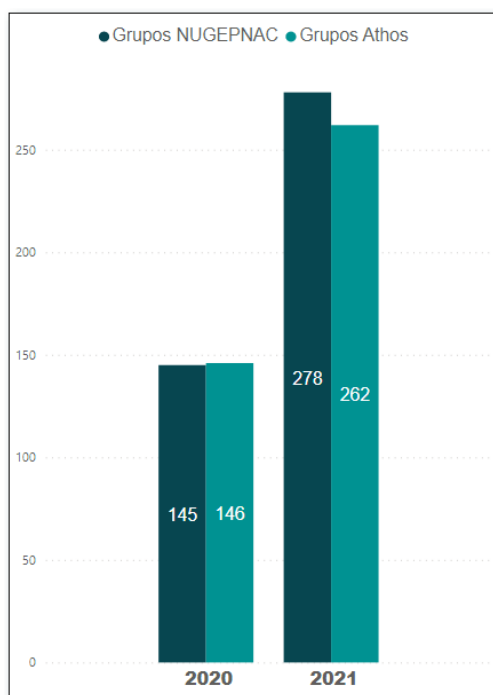
Assim, os números registrados em *log* indicam que somente em 2020 foram criados 146 grupos pelos operadores do sistema Athos. Os quais identificaram 1.694 processos paradigmas em um universo de 36.809 processos analisados. Já no ano de 2021, os registros indicam a criação de 262 grupos, isto é, 79,45% a mais que o ano anterior. A identificação de processos paradigmas também sofreu um salto de 34,06%, ao alcançar o número de 2.271.

<sup>50</sup> O log de transações registra todas as transações e as modificações de banco de dados feitas a cada requisição. Disponível em: < <https://docs.microsoft.com/pt-br/sql/relational-databases/logs/the-transaction-log-sql-server?view=sql-server-ver15>>. Acesso em: 20 dez. 2021.

O número de processos analisados, nesses dois critérios, em 2021 foi de 43.108. Ou seja, 17,11% a mais que em 2020.

Em um cenário comparativo entre os indicadores oficiais disponibilizados pela equipe do NUGEPNAC e os *logs* armazenados nos servidores da AIA, observamos que o Athos, no ano de 2020, criou 146 grupos, dos quais 144 foram utilizados pela equipe do Núcleo em seus estudos. No entanto, em 2021, a ferramenta de IA obteve um aproveitamento 79,45% superior aos grupos atualmente monitorados pelo Núcleo, ao disponibilizar 262 grupos, diante dos 278 em estudo no NUGEPNAC. O Gráfico 22 ilustra esta comparação.

Gráfico 22- Quantidade de grupos criados pelo Athos e NUGEPNAC entre os anos de 2020 e 2021.



Os impactos referentes ao destaque qualitativo imposto pelo rito dos repetitivos no âmbito do STJ, proporcionados pelos estudos realizados pela COGEPAC, com apoio da inteligência artificial têm sido impressionantes, pois a afetação e o julgamento de um recurso repetitivo no STJ refletem diretamente não apenas no âmbito do Poder Judiciário, mas também em toda sociedade.

Sanseverino e Marchiori (2020) trazem dois exemplos que ilustram a importância dessa atuação, sendo um da Seção de Direito Público (Primeira Seção) e outro da Seção de Direito Privado (Segunda Seção).

A Primeira Seção do STJ, com o julgamento do Tema Repetitivo n. 444, definiu o prazo prescricional para o redirecionamento da Execução Fiscal, estimando-se, conforme dados do Banco Nacional de Demandas Repetitivas e Precedentes Obrigatórios do Conselho Nacional de Justiça, que, na época da apreciação do tema, havia cerca de 12.000 processos sobrestados, aguardando a resolução da controvérsia. No entanto, o impacto desse tema chegou a seis milhões, somente no âmbito federal, conforme informações da Procuradoria-Geral da Fazenda Nacional.

Outro exemplo trazido pelos autores ilustra o impacto de um julgamento pelo STJ sob o rito dos repetitivos do Tema Repetitivo n. 710, em que se definiu a legalidade do sistema “*CreditScore*” desde que respeitados determinados requisitos. No momento da afetação da questão para julgamento repetitivo, identificou-se em torno de 200.000 ações somente no Estado do Rio Grande do Sul. Após o julgamento, apenas na Comarca de Florianópolis, foram julgados de uma só vez 55 mil processos por um único juiz, conforme informação do Tribunal de Justiça de Santa Catarina. Os autores concluem que a resolução qualificada da matéria pela Segunda Seção do STJ evitou o alastramento da questão para outros tribunais, pois, certamente, ela seria objeto de inúmeras ações pelo País em virtude do impacto direto nas relações comerciais de toda a sociedade, independentemente do local de domicílio (SANSEVERINO; MARCHIORI, 2020).

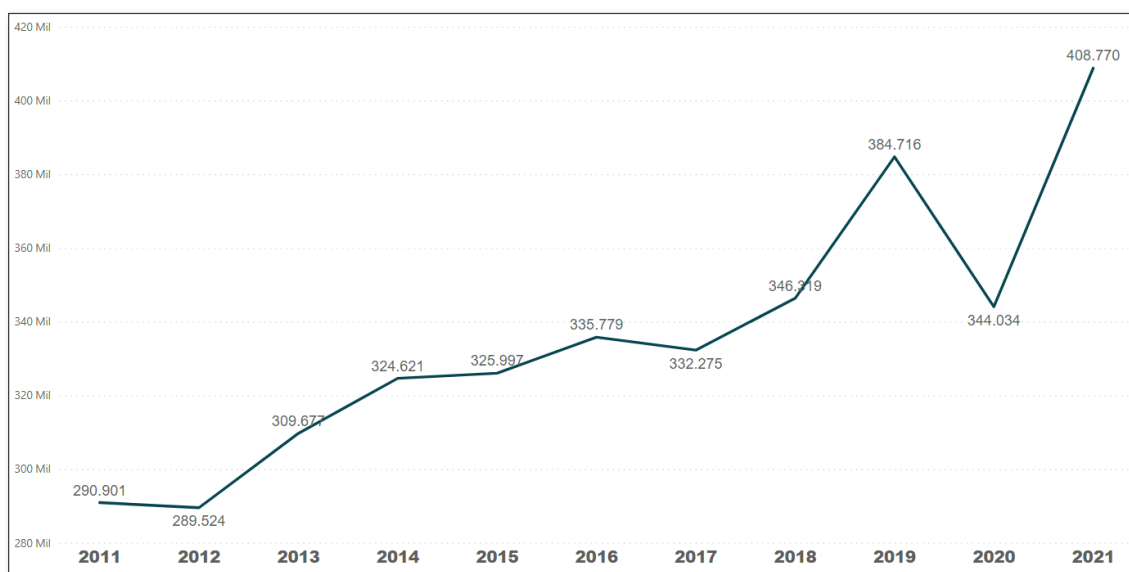
Ao reduzir a vazão de processos que ingressam no STJ, a sistemática dos recursos repetitivos alcança seu objetivo ao concretizar os princípios da celeridade na tramitação de processos, da isonomia de tratamento às partes processuais e da segurança jurídica<sup>51</sup>. Para tanto, outro impacto esperado, ao se estabelecer precedentes qualificados, será identificado no número de recursos recebidos que ingressam anualmente naquela Corte.

---

<sup>51</sup> Disponível em: <https://www.stj.jus.br/sites/portalp/Precedentes/informacoes-gerais/recursos-repetitivos>. Acesso em: 10 dez. 2021

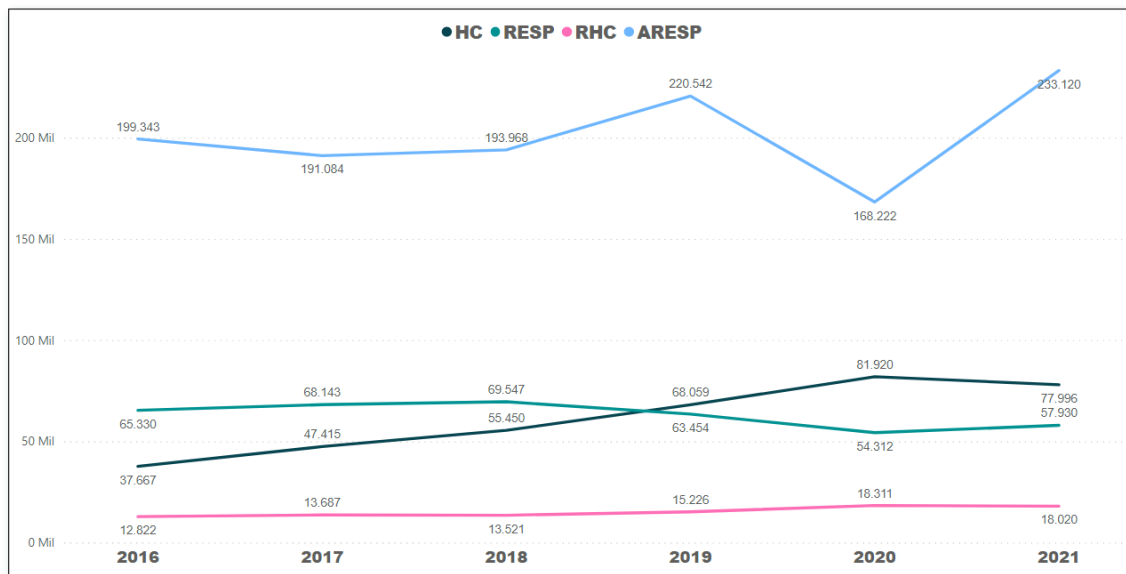
O boletim estatístico do STJ está em conformidade com seu Regimento Interno nos seus artigos 21, inciso XXIX, e 117, e publica mensalmente os dados quantitativos sobre o trâmite processual e a atividade judicante do Tribunal. Nele, encontramos dados estatísticos desde o ano 2000. Ao se efetuar a busca pelo número de processos recursais recebidos no STJ, nos últimos 10 anos, nos deparamos com um aumento de 40,51% neste número, onde partiu-se de 290.901 recursos em 2011, para 408.770 em 2021. Como visto no Gráfico 23:

Gráfico 23- Série histórica da quantidade de processos recebidos no STJ.



Porém, ao realizarmos um recorte dos principais recursos recebidos no STJ, a partir do ano de 2016, onde foi instituído o NUGEPNAC e iniciada a análise mais sistêmica dos Recursos Repetitivos imposta pelo CPC/2016. E, ao se incluir o ano de 2019, o qual deu-se início aos estudos mais criteriosos do rito dos repetitivos, com o auxílio da ferramenta de IA e a produção de indicadores de desempenho por parte dos operadores do Núcleo, nos deparamos com o cenário demonstrado no Gráfico 24:

Gráfico 24- Recorte dos principais recursos recebidos no STJ, a partir do ano de 2016.



O Gráfico 24 aponta o crescimento nos números de processos recebidos de praticamente todos principais recursos que ingressaram no STJ, com exceção dos Recursos Especiais. Como dito anteriormente, o REsp é o principal objeto de estudo da COGEPAC, com 98,27% dos Recursos Representativos da Controvérsia cadastrados na base de dados do STJ. De 2016 a 2021 houve redução de 11,32% na contagem de processos recebidos pelos indicadores estatísticos do Tribunal. Em sentido contrário, a contagem de AREsp, HC e RHC obtiveram aumento significativo em seus números. A Tabela 4 demonstra esta comparação:

Tabela 4- Redução de Recursos Especiais a partir do ano de 2016.

Recurso	Porcentagem
REsp	-11,32%
AREsp	16,95%
RHC	40,53%
HC	107,06%

Operadores do sistema Athos, tanto do NUGEPNAC, quanto da Secretaria Judiciária (Responsável pelo recebimento de todos os processos no STJ), ouvidos nesse estudo, informaram que ainda não é possível afirmar se somente o rito dos repetitivos, aliado ao sistema de IA, foram responsáveis pela

redução no número de Recursos Especiais ingressados anualmente naquela Corte. Isto porque, paralelamente, o Tribunal celebrou diversos acordos de cooperação técnica, com vistas ao avanço das ações de desjudicialização e redução de demandas com grandes litigantes. Tais acordos visam, principalmente, o mapeamento das controvérsias submetidas ao STJ e o monitoramento da atuação jurídica das partes, especialmente dos grandes demandantes, a fim de se executar estratégias voltadas a prevenção de litígios, gerenciamento de precedentes qualificados e de ações coletivas, e o fomento à resolução consensual das controvérsias<sup>52</sup>.

Destaca-se que, no referido projeto de redução de demandas, a solução de IA também produziu seus resultados, visto que, após o levantamento dos maiores litigantes do STJ, elaborado pela Secretaria Judiciária (SJD), o NUGEPNAC criou, por meio do Athos, os grupos para rastreamento das questões repetitivas dos maiores litigantes, bem como iniciou a análise destes grupos para fomentar a identificação de possíveis recursos representativos de controvérsias, e viabilizar a afetação de Temas para formação de Precedentes Qualificados. Foram criados 76 grupos de processos de demandas repetidas com mais de 50 processos cada, os quais também impactaram na redução de processos ingressados ao STJ.

Dados levantados pela Secretaria Judiciária, mostram que, até o final do mês de junho de 2021: 350.432 processos tiveram sua tramitação abreviada nas instâncias de origem (por desistência, acordo ou não interposição de recursos), ou seja, deixaram de ser enviados ao STJ. Houve a redução de aproximadamente 11,20% do número de AREsp's da AGU (total de 1.786 processos) em comparação com o período anterior equivalente; a taxa de decisões desfavoráveis foi reduzida em 14,15%; e foram homologadas 1.453 desistências no STJ<sup>53</sup>.

Por fim, a análise comparativa no recorte estatístico, do número de REsp's recebidos no STJ, entre os anos de 2016 e 2021, demonstra que até

---

<sup>52</sup> SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: processo SEI 028539/2020, acesso em: 15 dez. 2021

<sup>53</sup> SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: processo SEI 028539/2020. Acesso em: 15 dez. 2021

o ano de 2019, a redução no recebimento desse recurso foi de 2,87%, enquanto que, entre os anos de 2019 e 2021, a redução do mesmo quesito foi de 8,70%. Este indicador é de grande relevância, pois evidencia que os estudos proporcionados pela COGEPAC, aliados ao apoio da Inteligência Artificial, impactaram diretamente, e de forma significativa, no objetivo buscado há anos no Poder Judiciário: uma perspectiva de redução na vazão de demandas recebidas nos tribunais brasileiros.

## 5. Conclusão

Estudos apontam que a Inteligência Artificial passa por processos de amadurecimento. Relatórios como do Grupo Gartner identificam que a *hype* da IA está ficando para trás, e que tanto o mercado tecnológico, quanto às entidades responsáveis por estudos e pesquisas perceberam que a evolução dos sistemas em IA é lenta, complexa e cada vez mais distante do otimismo demonstrado no mercado de ficção científica.

O estado atual de conscientização direcionou novas pesquisas no sentido da qualidade dos dados utilizados nos sistemas de IA. Esta mudança de percepção fez com que os modelos de *Machine Learning* e Inteligência Artificial baseados em grandes quantidades de dados históricos se tornassem menos relevantes. Novos estudos indicam que a tendência será o uso dos dados disponíveis com maior eficiência, seja contando com dados amplos que permitem a análise e a convergência de uma variedade de fontes de dados, bem como dados específicos que capacitam a aplicação de técnicas analíticas que requerem menos dados.

Como verificado no presente estudo, os dados apontam o STJ alinhado à nova realidade consciente no uso e pesquisa da IA. Ao constatarmos a eficiência da IA na sistemática dos recursos repetitivos, no âmbito do STJ, concluímos que esta tecnologia se apresenta apenas como ferramenta de apoio, pois a real tomada de decisão, do que será um novo Tema Repetitivo, e seus reais impactos ao jurisdicionado, ainda se encontra fundamentada nas decisões da Comissão Gestora de Precedentes e de Ações Coletivas.



Porém, o sistema Athos permitiu uma mudança no sistema de avaliação de indicadores, daquela Corte. Durante muito tempo, o escopo avaliado foi a capacidade de vazão dos processos ingressados no STJ: comparando processos ingressados e processos baixados. A sistemática dos repetitivos exige avaliar o quanto a atuação do Tribunal consegue evitar a interposição de recursos, por meio da fixação de paradigmas consolidados. Para tanto, esse tipo de decisão estratégica depende do uso intensivo de dados: identificação de controvérsias repetitivas, segmentação da controvérsia em uma quantidade razoável de questões (garantindo que a decisão leve em conta a particularidade dos casos) e uma definição estratégica das teses a serem priorizadas para evitar a multiplicação de processos que seriam evitáveis por meio do estabelecimento imediato de padrões decisórios.

A adoção de estratégias de machine learning viabiliza ganhos relevantes na eficácia dos estudos produzidos pelo NUGEPNAC, o que possibilita aumento nos influxos processuais no STJ. No entanto, a necessidade de uma análise humana e política (que não é viável por IA) faz-se com que haja ainda a necessidade de um maior equilíbrio no fluxo de trabalho. Provavelmente, não será produtora o aumento de grupos de estudos no Núcleo, pois os operadores do NUGEPNAC não serão capazes de analisar a contento, e provavelmente, a capacidade de análise da COGEPAC é menor do que a capacidade da equipe técnica de apresentar propostas. Porém, chama atenção o desequilíbrio de *input* e *output* no sistema de produção de Temas repetitivos do STJ.

Os dados apresentados na presente pesquisa, indicam que o sistema Athos encontra-se em plena evolução, seus indicadores demonstram que tanto sua eficiência, quanto seus resultados estão em ascensão. Os indicadores que apontam a receptividade dos usuários diante da solução de Inteligência artificial foram positivos, ao constatarmos o aumento de 211,7% de requisições das funcionalidades disponíveis no sistema, entre os anos de 2020 e 2021. Assim, os dados sugerem que a solução de IA do STJ apresenta maturidade necessária para dar o aporte tecnológico necessário à possíveis mudanças de diretrizes da COGEPAC, no sentido de mitigar a divergência dos

indicadores de entrada e entrega do sistema de análise de novos Temas Repetitivos no STJ.

## 6. Referências

ABES- Associação Brasileira das Empresas de Software. Disponível em: <https://abessoftware.com.br/gartner-preve-que-o-mercado-global-de-inteligencia-artificial-chegara-a-us-625-bilhoes-em-2022/>. Acesso em: 05 fev. 2022.

AHIRE, Jayesh B. Introduction to Word Vectors. 2018. Disponível em <<https://medium.com/@jayeshbahire/introduction-to-word-vectors-ea1d4e4b84bf>>. Acesso em 20 dez. 2021.

BAHIA, Alexandre Gustavo Melo Franco; NUNES. Dierle José Coelho. Processo constitucional: uma abordagem a partir dos desafios do Estado Democrático de Direito. Revista Eletrônica de Direito Processual, v. 4, n. 4, 2009.

BECK, Ulrich. A metamorfose do mundo: novos conceitos para uma nova realidade. Trad. Maria Luiza X. de A. Borges. São Paulo: Zahar, 2018. p. 10-11.

BILLION, Arnaud; GUILLERMIN, Mathieu. Intelligence artificielle juridique: enjeux épistémiques et éthiques. Cahiers Droit, Sciences & Technologies, n 8. 2019, p131-147.

CHITTENDEM, Tara. Artificial Intelligence and the Legal Profession. Horizon Scan Forward Thinking, The Law Society, ago. 2017. Disponível em: [<https://bit.ly/2TYw3QN>]. Acesso em: 03 ago. 2021.

COELHO, Marcus Vinicius Furtado: Arts. 1.036 a 1.041 do CPC - Recursos repetitivos. Disponível em: <https://www.migalhas.com.br/coluna/cpc-marcado/307178/arts--1-036-a-1-041-do-cpc---recursos-repetitivos>. Acesso em: 02 set. 2021.

CONSELHO NACIONAL DE JUSTIÇA. Dados do Relatório Justiça em Números 2020, do. Justiça em Números 2020 ano base 2019/ Conselho Nacional de Justiça-Brasília: CNJ 2021.

CONSELHO NACIONAL DE JUSTIÇA. Disponível em: <https://www.cnj.jus.br/pesquisas-judiciarias/demandas-repetitivas/>. Acesso em: 01 jan. 2022.

CONSELHO NACIONAL DE JUSTIÇA. Justiça em Números: Disponível em: <<https://www.cnj.jus.br/pesquisas-judiciarias/justica-em-numeros/>>. Acesso em: 08 set. 2021.

CONSELHO NACIONAL DE JUSTIÇA: Painel de Consulta ao Banco de Dados Nacional de Demandas Repetitivas e Precedentes Obrigatórios. Disponível em: <[encurtador.com.br/AEKN5](http://encurtador.com.br/AEKN5)>. Acesso em: 03 fev. 2022.

CONSELHO NACIONAL DE JUSTIÇA: Painel de Consulta ao Banco de Dados Nacional de Demandas Repetitivas e Precedentes Obrigatórios. Disponível em: <[encurtador.com.br/alvBL](http://encurtador.com.br/alvBL)>.

CONSELHO NACIONAL DE JUSTIÇA: Painel de Consulta ao Banco de Dados Nacional de Demandas Repetitivas e Precedentes Obrigatórios. Disponível em: <[encurtador.com.br/gqvE2](http://encurtador.com.br/gqvE2) >. Acesso em: 20 jan. 2022.

CORMEN, Thomas H. Desmistificando Algoritmos. São Paulo: Elsevier, 2017. Dados (pp. 35-43). Lisboa: CEDIS. direito de não sujeição? In: F. P. Coutinho; G. C. Moniz, Anuário da Proteção de Dados.

CUEVA, Ricardo Villas Bôas, Inteligência Artificial no Judiciário. In: Inteligência Artificial e Direito Processual: Os impactos da virada Tecnológica no Direito Processual. Salvador: Editora JusPodivm, 2021.

DEPARTMENT OF COMPUTER ENGINEERING, BOĞAZIÇI UNIVERSITY. <https://www.cmpe.boun.edu.tr/~akin/cmpe223/chap1.htm>. Acesso em 30 de julho de 2021.

ENGELMANN, Wilson; WERNER, Deivid Augusto. Inteligência Artificial e Direito. In: Inteligência Artificial e Direito: ética, regulação e responsabilidade. 1. Ed. São Paulo: Thomson Reuters Brasil, 2021.

GARTNER. Disponível em: <https://www.gartner.com/en/topics/artificial-intelligence>. Acesso em 23 dez. 21.

GLAVAS, F. G., KARAN, M., SNAJDER, J. & BASIC, B. D. "TakeLab: Systems for Measuring Semantic Text Similarity". in SEM 2012: The First Joint Conference on Lexical and Computational Semantics, 2012.

GUTIERREZ, Andriei. É possível confiar em um sistema de Inteligência Artificial? Práticas em torno da melhoria da sua confiança, segurança e evidências de accountability. In: Inteligência Artificial e Direito: ética, regulação e responsabilidade. 1. Ed. São Paulo: Thomson Reuters Brasil, 2020.

HENARCOS. Disponível em: <https://henarcos.com.br/depois-hype-inteligencia-artificial/>. Acesso em 21 jan. 2022.

IANSTITI, Marco; LAKHANI, Karim R. A era da Inteligência Artificial. 1.Ed. Cascavel/PR. Editora Alfacon, 2021.

INFORCHANNEL. Disponível em: <https://inforchannel.com.br/2021/03/29/gartner-identifica-as-10-principais-tendencias-em-data-analytics-para-2021/>>. Acesso em: 07 fev. 2022.

KLAUS Schwa. A Quarta Revolução Industrial. 1. Ed. São Paulo: Editora Edi, 2018.

LAGE, Fernanda de Carvalho. Manual de Inteligência Artificial no Direito Brasileiro. 1 Ed. Salvador: Editora JusPodivm, 2021.

LE, Quoc V. MIKOLOV, Tomas: “Distributed Representations of Sentences and Documents”, Mountain View, CA: Google Inc. 2014.

MANCUSO, Rodolfo de Camargo. Recurso Extraordinário e Recurso Especial. 2. Ed.rev. atual. e ampl. de acordo com o vidente CPC. São Paulo: Thomson Reuters Brasil, 2018.

MANNING, C. & SCHÜTZE, H. “Foundations of Statistical Natural Language Processing”. MIT Press, 2000.

MARTINS, Amilar Domingos Moreira. Agrupamento Automático de Documentos Jurídicos com uso de Inteligência Artificial. 1 Ed. Brasília: IDP, 2018.

MICROSOFT INC. Disponível em: < <https://docs.microsoft.com/pt-br/sql/relational-databases/logs/the-transaction-log-sql-server?view=sql-server-ver15>>. Acesso em: 20 dez. 2021.

MULHOLLAND, Caitlin. Inteligência Artificial e Direito: ética, regulação e responsabilidade. 2. Ed. São Paulo: Thomson Reuters Brasil, 2020.

NORONHA, João Otávio de. Os avanços da inteligência artificial no STJ: caminho para uma justiça mais rápida e efetiva. In: O Direito Civil na era da Inteligência Artificial [livro eletrônico] / Rodrigo da Guia Silva e Gustavo Tepedino coordenadores. 1. Ed. São Paulo: Thomson Reuters Brasil, 2020.

NUNES, Marcelo Guedes. Jurimetria: Como a Estatística pode reinventar o direito. 1 Ed. São Paulo: Editora Revista dos Tribunais, 2016.

O'NEIL, Cathy. Weapons of Math Destruction: How de Big Data increases inequality and threatens democracy. 1. Ed. Nova Iorque: Crown Publishers, 2016.

PEIXOTO, Fabiano Hartmann; SILVA, Roberta Zumplick Martins da. Inteligência Artificial e Direito: 1. ed. Curitiba: Alteridade, 2019.

POLIDO, Fabrício Bertini Pasquot. Novas perspectivas para regulação da IA: Diálogos entre as políticas domésticas e os processos transnacionais. In: Inteligência Artificial e Direito: ética, regulação e responsabilidade. 2. Ed. São Paulo: Thomson Reuters Brasil, 2020.

ROCKCONTENT. Disponível em <https://rockcontent.com/br/blog/quadrante-magico-gartner/>. Acesso em: 23 dez. 2021.

RODRIGUES, Bruno Alves. A inteligência artificial no poder judiciário [livro eletrônico]: e a convergência com a consciência humana para a efetividade da justiça / Bruno Alves Rodrigues. 1. ed. São Paulo: Thomson Reuters Brasil, 2021.

ROUGHGARDEN, Tim. Algorithms Illuminated: Part 1. São Francisco: Soundli Keyyourself Publishing, 2017. Disponível em: <http://algorithmsilluminated.org/>. Acesso em 20 dez. 2021.

SANSEVERINO, Paulo de Tarso; MARCHIORI, Marcelo Ornelas. O projeto Athos de inteligência artificial e o impacto na formação de precedentes qualificados no Superior Tribunal de Justiça. In: O Direito Civil na era da Inteligência Artificial [livro eletrônico] / Rodrigo da Guia Silva e Gustavo Tepedino coordenadores. 1. ed. São Paulo: Thomson Reuters Brasil, 2020.

SANTOS, Silas Silva; CUNHA, Fernando Antônio Mais da; FILHO, Milton Paulo de Carvalho; RIGOLIN, Antônio. Comentários ao Código de Processo Cível:

perspectiva da magistratura/ coordenação Silas Silva Santos [et al.]. -- 2. ed. São Paulo: Thomson Reuters Brasil, 2020.

SARTOR, G.; BRANTING, L. Karl. Introduction: judicial applications of Artificial Intelligence. *Artificial Intelligence and Law*, v. 6, p. 105-110, 1998. Disponível em: <<https://bit.ly/2G6zyQz>>. Acesso em: 20 DEZ. 2021.

SARTOR, G.; BRANTING, L. Karl. Introduction: judicial applications of Artificial Intelligence. *Artificial Intelligence and Law*, v. 6, p. 105-110, 1998. Disponível em: [<https://bit.ly/2G6zyQz>]. Acesso em: 09 ago. 2021.

SERVIÇO BRASILEIRO DE APOIO ÀS MICRO E PEQUENAS EMPRESAS. Disponível em: [www.sebrae.com.br](http://www.sebrae.com.br). Acesso em: 09. ago. 2021.

SILVA, Nilton Correia. Compreensão da Inteligência Artificial e dos seus Pressupostos de Controle e Regulação. In: *Inteligência Artificial e Direito: ética, regulação e responsabilidade*. 2. Ed. São Paulo: Thomson Reuters Brasil, 2020.

SINCLAIR, J. 2005. "Corpus and Text - Basic Principles in Developing Linguistic Corpora: a Guide to Good Practice". ed. M. Wynne. Oxford: Oxbow Books: 1-16. Disponível em <https://ota.ox.ac.uk/documents/creating/dlc/chapter1.htm>. Acesso em 20 dez. 2021.

SMARTCONSULTING. Disponível em: <<https://blog.smartconsulting.com.br/quadrante-magico-gartner/>>. Acesso em 27 dez. 21.

SOURDIN, Tania. Judge v Robot? Artificial Intelligence and judicial decision-making. *University of New South Wales Law Journal*, v. 41, n. 4, p. 1114-1133, 2018. Disponível em: [<https://bit.ly/2D2HCAP>]. Acesso em: 03 ago. 2021.

SOUZA, Carlos Affonso Pereira de; OLIVEIRA, Jordan Vinícius de. Sobre os ombros de robôs? A Inteligência Artificial entre fascínios e desilusões. In: *Inteligência Artificial e Direito: ética, regulação e responsabilidade*. 1. Ed. São Paulo: Thomson Reuters Brasil, 2020.

STEIBEL, Fabro; VICENTE, Victor Freitas; JESUS, Diego Santos Vieira de. Possibilidade e potenciais da utilização da Inteligência Artificial. In: *Inteligência Artificial e Direito: ética, regulação e responsabilidade*. 2. Ed. São Paulo: Thomson Reuters Brasil, 2020.

SULOCKI, Victoria de. Novas tecnologias, velhas discriminações: ou da falta de reflexão sobre o sistema de algoritmos na Justiça Criminal. In: Inteligência Artificial e Direito: ética, regulação e responsabilidade. 1. Ed. São Paulo: Thomson Reuters Brasil, 2020.

SUPERIOR TRIBUNAL DE JUSTIÇA Disponível em: <[https://processo.stj.jus.br/repetitivos/temas\\_repetitivos/](https://processo.stj.jus.br/repetitivos/temas_repetitivos/)>. Acesso em 20 jan. 2022.

SUPERIOR TRIBUNAL DE JUSTIÇA. Boletim Estatístico de dezembro de 2021. Disponível em: <<https://www.stj.jus.br/webstj/Processo/Boletim/sumario.asp>>. Acesso em: 15 jan. 2022.

SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: <[https://processo.stj.jus.br/repetitivos/temas\\_repetitivos/](https://processo.stj.jus.br/repetitivos/temas_repetitivos/)>. Acesso em: 10 set. 2021.

SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: <<https://www.stj.jus.br/sites/portalp/Precedentes/informacoes-gerais/recursos-repetitivos>>. Acesso em: 06 set. 2021.

SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: <<https://www.stj.jus.br/sites/portalp/Precedentes/informacoes-gerais/controversias>>. Acesso em: 23 de nov. 2021.

SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: <https://bityli.com/cq0pg>. Acesso em: 10 set. 2021.

SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: <https://www.stj.jus.br/sites/portalp/Institucional/Historia/A-crise-do-Supremo>. Acesso em: 08 set. 2021.

SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: <https://www.stj.jus.br/sites/portalp/Inicio>. Acesso em 20 jan. 2022.

SUPERIOR TRIBUNAL DE JUSTIÇA. Disponível em: processo SEI 028539/2020, acesso em: 15 dez. 2021.



SUPERIOR TRIBUNAL DE JUSTIÇA: Disponível em: <<https://www.stj.jus.br/sites/portalp/Precedentes/informacoes-gerais/comissao-gestora-de-precedentes-e-de-acoes-coletivas>>. Acesso em: 03 dez. 2021.

SUPERIOR TRIBUNAL DE JUSTIÇA: Disponível em: <https://www.stj.jus.br/sites/portalp/Precedentes/informacoes-gerais/controversias>. Acesso em: 20 dez. 2021.

SUPERIOR TRIBUNAL DE JUSTIÇA: Relatório de encerramento projeto Athos. Rev. 06.03.

SUPREMO TRIBUNAL FEDERAL. Disponível em: < <https://bityli.com/Tuxy1>>. Acesso em: 08 set 2021.

SYAM, Niladri; SHARMA, Arun. Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice. *Industrial Marketing Management*, n. 69, p. 135-146, 2018.

TEIXEIRA, Tarcisio; CHELIGA, Vinicius. *Inteligência Artificial: Aspectos Jurídicos*. 3.Ed. Salvador: Editora JusPodivm, 2021.

VERONESE, Alexandre; SILVEIRA, Alessandra; Lemos, Amanda Nunes Lopes Espiñeira Lemos. *Inteligência Artificial, mercado único digital e a postulação de um direito às inferências justas e razoáveis: uma questão jurídica entre a ética e a técnica*. In: *Inteligência Artificial e Direito: ética, regulação e responsabilidade*. 2. Ed. São Paulo: Thomson Reuters Brasil, 2020.

VIEIRA, Renata; LOPES, Lucelene. *Processamento de Linguagem Natural e o Tratamento Computacional de Linguagens Científicas*. in PERNA et al. *Linguagens Especializadas em Corpora - Modos de Dizer e Interfaces de Pesquisa*. Porto Alegre: EDIPUCRS, 2010.