

## Aplicação da Inteligência Artificial na identificação de conexões pelo fato e tese jurídica nas petições iniciais e integração com o Sistema de Processo Eletrônico

Antônio Pires de Castro Júnior  
Wesley Pacheco Calixto  
Cláudio Henrique Araujo de Castro

Resumo: Esse artigo trabalha com a possibilidade de identificar e unificar, automaticamente, volumes significativos de demandas judiciais em tramitação que possuam o mesmo fato e tese jurídica. Com a identificação e a unificação dos processos em agrupamentos, objetiva-se criar pendências no Sistema de Processo Eletrônico com a finalidade de informar a possibilidade de ocorrência de conexão às diferentes unidades judiciais que receberam as causas por distribuição, alertando e facilitando a análise pelo Julgador. São aplicadas técnicas de Processamento de Linguagem Natural, aprendizagem por similaridade e Redes Neurais Artificiais. A solução de Inteligência Artificial (IA) construída, chamada Berna, encontra-se em produção no Poder Judiciário Goiano. A precisão de 96% nos estudos de casos demonstra a efetividade do método.

Palavras-chaves: Processamento de linguagem natural; aprendizagem por similaridade; redes neurais artificiais; petição inicial; sistema de processo eletrônico.

Abstract: This article works with the possibility to automatically identify and unify big data of lawsuits in progress that have the same fact and thesis. With the identification and unification of the processes in *clusters*, the objective is to create disputes in the Electronic System of Lawsuits, in order to inform the possibility of connection with the different judicial units that received the cases by distribution, alerting and facilitating the analysis by the Judge. Techniques of Natural Language Processing, similarity learning and Artificial Neural Networks are applied. The Artificial Intelligence solution built, called Berna, is in production mode in the Judiciary of the State of Goiás. The 96% accuracy in case studies demonstrates the effectiveness of the method.

Keywords: Natural language processing; similarity learning; artificial neural networks; judicial complaint; electronic system of lawsuits.

### 1. Introdução

No relatório Justiça em Números 2019 (KIM *et al*, 2019), percebe-se que o número de processos pendentes na Justiça Estadual vem aumentando a cada ano. Esse indicador crescente sinaliza situações de superlotação nas unidades judiciárias, que a cada dia veem o seu quadro de servidores e magistrados diminuindo, prejudicando o rápido atendimento aos anseios e direitos da sociedade.

Dados estatísticos reunidos pelo Conselho Nacional de Justiça, como o informado acima, permitem aplicar importantes inferências e levantar hipóteses de situações que possam acontecer no Judiciário, como: i) há demandas idênticas tramitando na justiça em unidades judiciais diferentes, que maculam o princípio do Juiz natural; e, ii) nas várias demandas em tramitação, existe número expressivo de processos distribuídos para unidades judiciais distintas que contém na sua petição inicial o mesmo fato gerador e a mesma tese jurídica, os quais não são identificados pelos atuais mecanismos de detecção de conexão nos sistemas eletrônicos de processos.

Nas duas hipóteses levantadas acima, as análises para tentar identificar não seriam as mesmas comumente aplicadas hoje nos sistemas de andamento processual e/ou de processo eletrônico, que checam nome das partes, CPFs/CNPJs, classes, assuntos, endereços ou outros metadados cadastrados, procurando em campos nos bancos de dados para tentar encontrar semelhanças. Este artigo propõe construir buscas que utilizam tecnologias de inteligência artificial, que possam analisar, checar e comprovar essas hipóteses nos textos dos documentos, mais precisamente mediante checagem do inteiro teor das petições iniciais.

Ora, é conhecido e notório que grande parte das petições iniciais nas ações judiciais em curso seguem modelos disponíveis na internet, sendo que alguns demandantes têm utilizado, inclusive, *softwares* de inteligência artificial para gerá-las automaticamente.

Assim, buscar por inovações em ferramentas de apoio ao julgador, objetivando promover o aperfeiçoamento da gestão do conhecimento no judiciário (CASTRO; CALIXTO; FRANCO, 2014) constitui-se ação importante e premente na atualidade.

O objetivo deste artigo é apresentar e aplicar modelo de inteligência artificial que possibilite identificar e verificar se as hipóteses descritas acima são verdadeiras. Nesse sentido, este trabalho propõe-se a apresentar método e ferramenta tecnológica que permita minerar as petições iniciais dos processos em tramitação, identificando duas situações: i) processos diferentes com petições iniciais idênticas em todos os aspectos e ii) processos com o mesmo fato e tese jurídica.

Ainda, caso sejam comprovadas as hipóteses, tem-se também como objetivo que a ferramenta possa incluir pendências no Sistema de Processo Eletrônico, para que o julgador possa, ao tomar conhecimento das similaridades, analisar e decidir ou julgar conforme seu entendimento.

Este documento contém a seguinte estrutura: na seção 2, descreve o desenvolvimento do trabalho, que contém breve motivação, metodologia desenvolvida e resultados aplicados pela ferramenta construída em ambiente de produção. Na seção 3, informa as conclusões do artigo e, na seção 4, as referências bibliográficas.

## 2. Desenvolvimento

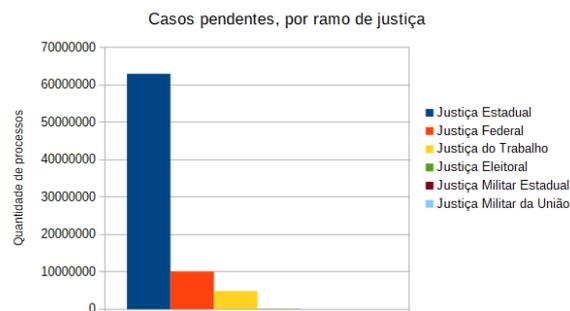
É natural que o Poder Judiciário acompanhe o progresso tecnológico e adote novas formas de gestão de informação por meio de sistemas que concretizem a aplicação dos princípios constitucionais e processuais, como o da eficiência e da celeridade processual (SCHAEFER, 2005).

A eficiência é princípio da Administração Pública (TEMER; MAGALHAES, 1998), sendo o Judiciário o braço forte no estabelecimento da paz e da justiça entre os poderes e a sociedade em geral, tendo como obrigação garantir qualidade, dinamicidade e celeridade na prestação dos serviços públicos (RAMOS; ROVER, 2005).

### 2.1 Motivação: Justiça em Números

Uma das grandes inovações no Judiciário Brasileiro e idealizado pelo Conselho Nacional de Justiça é a apresentação de dados estatísticos anuais para todas as esferas da Justiça, compilados no tradicional e reconhecido “Justiça em Números” (KIM *et al*, 2019). Segundo esse caderno estatístico, o Judiciário no Brasil encerra 2018 com quase 80 milhões de processos em andamento (KIM *et al*, 2018). A figura 1 mostra os números por segmento de justiça.

Figura 1 – Número de processos pendentes por ramo de justiça



Fonte: caderno “Justiça em Números 2019” (KIM *et al*, 2019).

Outro dado relevante é a série histórica dos casos pendentes na Justiça Estadual, percebe-se que esse número não diminui com os anos (KIM *et al*, 2019), figura 2. Esses números são os reflexos do aumento de casos novos e dos casos reativados.

Figura 2 – Série histórica dos processos pendentes na Justiça Estadual



Fonte: caderno Justiça em Números 2019 (KIM *et al*, 2019).

Esses números justificam a realização de estudos envolvendo tecnologia da informação para tentar ajudar e/ou facilitar a análise e julgamento dos processos judiciais. Essa é a principal motivação deste trabalho.

### 2.2 Metodologia

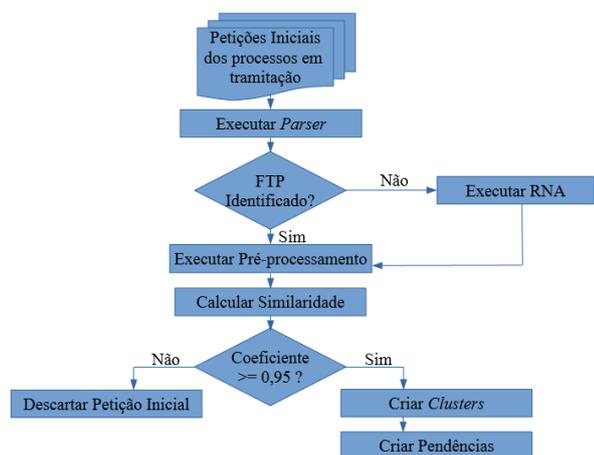
Visto que os estudos são aplicados na área do Direito, é relevante informar que outras áreas do conhecimento também têm aplicado e trabalhado na classificação de documentos, como: i) Medicina (ARSENE; DUMITRACHE; MIHU, 2011), ii) Biologia (LAMY, 2017), iii) Engenharia (RANI; DHAR; VYAS, 2017), iv) Direito (ZHANG *et al*, 2017), v) Educação (GRUBISIC; STANKOV; PERAIC, 2013), entre outros.

Estudos relacionados podem ser acompanhados em: i) CECI; GANGEMI, 2016; ii) FAWEL *et al*, 2015; iii) CALAMBAS *et al*, 2015; iv) ZHANG; WANG; PU, 2015; v) ZHANG *et al*, 2017 continuam o trabalho em ZHANG; WANG; PU, 2015 e vi) RANI; DHAR; VYAS, 2017.

Sabendo antecipadamente de que os textos judiciais eletrônicos para análise e processamento são volumosos e em grande quantidade para o projeto em tela, ainda que não estejam estruturados em modelos padronizados, faz-se necessário aplicar técnicas computacionais de Processamento de Linguagem Natural (PLN). O PLN é a área da inteligência artificial que possibilita aos computadores analisar, manipular e interpretar a linguagem humana.

Para que seja possível manipular o grande volume de textos judiciais, objetivando criar *clusters* de similares, o método proposto pode ser entendido pelo fluxograma da figura 3, na qual os procedimentos estão divididos em fases.

Figura 3 – Método proposto para identificação de petições iniciais similares em atendimento à hipótese 2. FTP significa: Fato, Tese e Pedido



Fonte: os autores (2020).

A figura 3 apresenta o método proposto em passos objetivando atender à hipótese 2 na introdução. Para atendimento à hipótese 1, não são necessários os passos Executar *Parser* e Executar RNA bem como o passo de decisão FTP Identificado. Ainda, para hipótese 1, o passo de decisão Coeficiente  $\geq 0,95$  deve ser substituído pelo passo de decisão Coeficiente = 1.

### 2.2.1 Separar os documentos em partes (fase, tese e pedido)

Em que pese não seja necessário para a primeira hipótese, porém para a segunda hipótese faz-se necessário identificar e separar os trechos nas petições iniciais em 3 (três) partes: i) parte que conta a história geradora do fato; ii) parte que traz o embasamento jurídico para entrar com a ação; e iii) parte que o demandante/requerente/autor faz a solicitação ao Poder Judiciário, sendo essas partes doravante denominadas: i) Fato

jurídico; ii) Tese jurídica e iii) Pedido (FTP). Para conseguir cumprir, são aplicados dois modelos, excludentes entre si:

- Parser: software* que consegue identificar e separar o inteiro teor do documento em pedaços. O *Parser* é construído com palavras conhecidas que permitem dividir o texto sempre que encontrar essas palavras comuns.
- Redes Neurais Artificiais *Perceptron* (RNA): As RNAs são utilizadas para fazerem previsões dos trechos, frases, no inteiro teor dos documentos.

Como apresentado na figura 3, na primeira análise é utilizado o *Parser* com palavras previamente conhecidas para conseguir separar o FTP. A figura 4 demonstra alguns termos conhecidos utilizados pelo método. Caso não seja possível realizar a separação dos trechos pelo *Parser*, aplica-se, então, a RNA.

Figura 4 – Tela de administração para inclusão, edição e exclusão de *Parsers*

### Listagem: Parser

Registrar Tela de Listagem de Documentos Indexados

Tipo	Contexto	Ativo			
Pedido	Isto posto	SIM	Mostrar	Editar	Excluir
Pedido	Ante o exposto	SIM	Mostrar	Editar	Excluir
Fato	Do fato	SIM	Mostrar	Editar	Excluir
Fato	Dos fatos	SIM	Mostrar	Editar	Excluir
Tese	Do direito	SIM	Mostrar	Editar	Excluir
Tese	Dos direitos	SIM	Mostrar	Editar	Excluir
Pedido	Ex positis	SIM	Mostrar	Editar	Excluir

Fonte: os autores (2020).

As RNAs são utilizadas para fazerem previsões na análise de textos nos documentos. A RNA utilizada é uma rede neural *perceptron* com duas camadas e treinamento supervisionado, tendo a função de ativação o arco tangente hiperbólico, em que o valor de saída da rede fica entre  $[-1,1]$ . Essa RNA só é utilizada se o *Parser* para encontrar os pedaços (fato, tese e pedido) no inteiro teor do texto falhar. Para treinamento dessa rede, foram extraídos, numa amostra de 100 documentos, os *tokens* de maior frequência dentro dos fatos jurídicos, das teses jurídicas e dos pedidos.

O modelo de RNA é utilizado para fazer a previsão dos trechos das petições iniciais. Para isso, faz-se necessário separar o texto em frases, depois separar a frase em

palavras. As palavras são utilizadas como entrada da RNA, sendo que na saída dessa rede neural temos números entre  $[-1,1]$ , que nos permitem prever conforme explicado na tabela 1.

Tabela 1 – Indica os valores de saída da RNA *perceptron* e os seus significados de predição

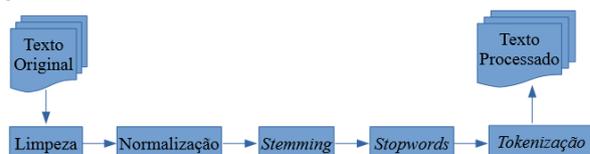
Valor de saída da Rede	Predição da RNA
$0,5 < \text{valor} \leq 1$	Identifica que faz parte do fato jurídico.
$0 < \text{valor} \leq 0,5$	Identifica que faz parte da tese jurídica.
$-0,5 < \text{valor} \leq 0$	Identifica que faz parte do pedido.
$-1 < \text{valor} \leq -0,5$	Não consegue prever.

Fonte: os autores (2020).

## 2.2.2 Pré-processamento

Após a etapa de separar o FTP, parte-se para a aplicação dos métodos de pré-processamento, geralmente, utilizados em PLN para cada um dos textos judiciais recebidos, conforme sequência da Figura 5.

Figura 5 – Pré-processamento PLN para os textos judiciais recebidos



Fonte: os autores (2020).

- limpeza – retirar caracteres especiais, pontuações, parênteses, hifens e outros caracteres sem valor no contexto;
- normalização – retirar os acentos, cedilha e colocar todas as palavras em caixa-baixa;
- stemming* – reduzir palavras flexionadas ou derivadas da sua base;
- stopwords* – retirar os artigos, preposições, conjunções sem significados, palavras que não tenham sentido semântico ao texto, incluem neste pré-processamento palavras comuns na área do Direito sem importância para o contexto;
- tokenização* – transformar o texto em vetor de palavras, retirando suas repetições.

O pré-processamento é feito tanto no inteiro teor do texto original, quanto nas partes FTP encontradas. Após esse passo de pré-processamento, o texto tratado é gerado e gravado, facilitando a análise da próxima etapa.

## 2.2.3 Cálculo de Similaridade

Após as etapas de 2.2.1 e 2.2.2 expostas acima — que criam dois conjuntos de documentos: i) documentos em que foram separados o fato, a tese e o pedido (FTP) e ii) documentos que não foram separados o FTP —, parte-se para a etapa de maior importância do método proposto.

A presente etapa consiste em utilizar técnicas de Recuperação de Informações para calcular a similaridade, objetivando identificar documentos idênticos ou parecidos e separá-los em *clusters* de similares, tanto para os documentos em que foi possível separar os pedados (fato, tese e pedido), quanto para os documentos nos quais não foi possível realizar a separação.

Na área de recuperação de informações existem várias técnicas utilizadas para buscar documentos, fazendo o cálculo da expressão de busca com cada documento no *corpus*. Nessa busca podem ser utilizadas várias técnicas, como: i) lógica, ii) estatística, iii) teoria de conjuntos e iv) outras. Boughanem, Brini e Dubois (2009) afirmam que modelos estatísticos têm impulsionado o desenvolvimento de novas soluções em sistemas de recuperação textual, incluindo: i) booleanos, ii) vetor, iii) probabilístico e iv) agrupamento. A eficiência do sistema de RI é diretamente vinculada ao modelo aplicado.

Thada e Jaglan (2013) informam alguns modelos tradicionais para calcular a similaridade, como: i) Jaccard; ii) Cosseno e iii) coeficiente de Dice, sendo todos os modelos estatísticos que geram seus resultados entre  $[0,1]$ .

Nos estudos de Thada e Jaglan (2013) percebe-se que o modelo de Jaccard apresenta os menores valores estatísticos de comparação, informando menores distâncias entre os modelos comparados. No trabalho de Thada e Jaglan (2013), por Jaccard apresentar o menor percentual estatístico diante dos modelos Cosseno e Dice, entende-se que coeficiente entre  $[0,1]$  informado pela expressão de Jaccard terá um grau de similaridade proporcional superior aos outros dois modelos. Todos esses modelos são indicados para o processamento de grandes volumes de dados, mas como o objetivo deste trabalho é encontrar documentos com alto grau de similaridade, optou-se por utilizar o modelo de Jaccard.

A expressão de Jaccard mede a similaridade entre conjuntos de amostras finitas e é definido como o tamanho da interseção dividido pelo tamanho da união dos conjuntos de amostras. Dada a expressão em (1), Jaccard mede a similaridade entre documentos  $D_1$  e  $D_2$ .

$$Jaccard(D_1, D_2) = \frac{|V_{D_1 \cap V_{D_2}}|}{|V_{D_1 \cup V_{D_2}}|} = \frac{|V_{D_1 \cap V_{D_2}}|}{|V_{D_1}| + |V_{D_2}| - |V_{D_1 \cap V_{D_2}}|} \quad (1)$$

em que  $V_{D_1}$  e  $V_{D_2}$  são vetores que contêm os *tokens* dos seus documentos  $D_1$  e  $D_2$ .

Para a primeira hipótese, na representação em (1), considera-se  $D_1$  e  $D_2$  o inteiro teor dos documentos. Já na segunda hipótese, na representação em (1), considera-se  $D_1$  os trechos fato jurídico e tese jurídica do documento  $D_1$  bem como considera-se  $D_2$  os trechos fato jurídico e tese jurídica do documento  $D_2$ .

Pelo resultado do cálculo de similaridade entre o fato jurídico e tese jurídica nos documentos  $D_1$  e  $D_2$ , o método poderá, então, deduzir as situações:

- Hipótese 1: se o resultado for igual a 1, significa que  $D_1$  é igual a  $D_2$ , ou seja, existem petições iniciais idênticas em processos diferentes;
- Hipótese 2: se o resultado for igual ou maior que 0,95, pode-se afirmar que trata-se de mesmo fato e tese jurídica em processos diferentes.

### 2.3 Resultados

A aplicação computacional do método apresentado em 2.2 resulta na ferramenta de Inteligência Artificial chamada Berna. Essa ferramenta recebe como entrada as petições iniciais de vários processos em tramitação e constroem *clusters* para os casos em que o cálculo de similaridade no inteiro teor dos documentos é igual a 1 e para os casos em que o cálculo de similaridade entre o fato e tese jurídicas são maiores e iguais a 0,95.

Os estudos de casos foram realizados no Poder Judiciário do estado de Goiás, em dados reais, nos processos em tramitação. As unidades foram:

- 11 Juizados Especiais Cíveis na comarca de Goiânia, entrância final;
- 1ª Vara da Fazenda Pública Municipal na comarca de Goiânia, entrância final;
- Nas Turmas Recursais;
- E na Comarca de Pontalina de Goiás, entrância intermediária.

A tabela 2 informa a quantidade de petições iniciais de processos pendentes que foram importadas para a Berna, ou seja, utilizados nos estudos.

Tabela 2 – Relação das unidades em que o estudo de caso foi aplicado e a quantidade de documentos analisados/processados

Unidades Judiciais	Petições Iniciais
Goiânia – 1º Juizado Especial Cível	3.474
Goiânia – 2º Juizado Especial Cível	3.347
Goiânia – 3º Juizado Especial Cível	3.517
Goiânia – 4º Juizado Especial Cível	1.961
Goiânia – 5º Juizado Especial Cível	2.745
Goiânia – 6º Juizado Especial Cível	3.672
Goiânia – 7º Juizado Especial Cível	4.071
Goiânia – 8º Juizado Especial Cível	3.205
Goiânia – 9º Juizado Especial Cível	3.886
Goiânia – 10º Juizado Especial Cível	3.338
Goiânia – 11º Juizado Especial Cível	3.156
Goiânia – 1ª Vara da Fazenda Pública Municipal	145.691
1ª Turma Recursal dos Juizados Especiais	5.435
2ª Turma Recursal dos Juizados Especiais	6.072
3ª Turma Recursal dos Juizados Especiais	7.034
4ª Turma Recursal dos Juizados Especiais	7.885
Pontalina – 1º Juizado Especial Cível	101
Pontalina – 2º Juizado Especial Cível	136
Pontalina – 1ª Vara Cível	179
Pontalina – 2ª Vara Cível	270
Pontalina – Juizado Especial Criminal	6
Pontalina – Vara das Fazendas Públicas	1.173
Pontalina – Vara de Família e Sucessões	105
Total	210.459

Fonte: os autores (2020).

Desses números, chama a atenção a quantidade de processos analisados pela Berna na 1ª Vara da Fazenda Pública Municipal de Goiânia.

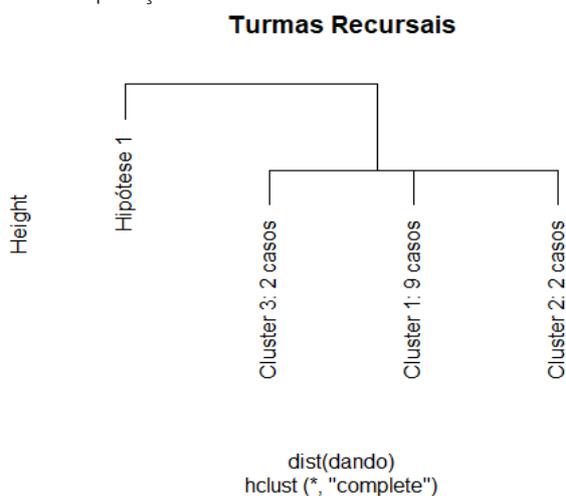
#### 2.3.1 Casos de 100% de similaridade no inteiro teor da petição inicial

Na análise da primeira hipótese, os estudos se concentraram nas petições iniciais nos processos pendentes das Turmas Recursais e na 1ª Vara da Fazenda Pública Municipal de Goiânia. São apresentados os resultados de casos reais encontrados pela Berna que têm a mesma petição inicial protocolada no Judiciário, gerando processos com números diferentes.

Com relação às Turmas Recursais, a metodologia aplicada pela Berna conseguiu encontrar 13 petições iniciais idênticas, protocoladas em diferentes processos, distribuídas nos *clusters* conforme demonstra-

do no Dendograma da figura 6. Percebe-se que no *cluster* 1 foram identificados 9 casos de petições idênticas.

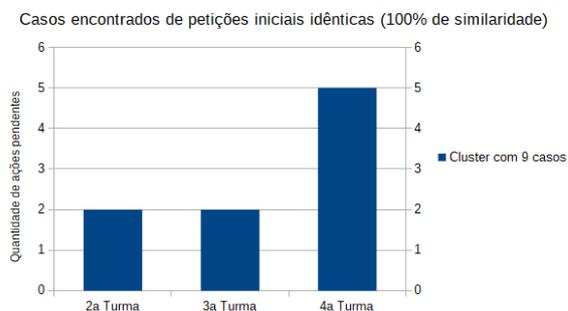
Figura 6 – Representação dos *clusters* encontrados nos casos de similaridade 100% no inteiro teor da petição inicial nas Turmas Recursais



Fonte: os autores (2020).

No gráfico da figura 7, é possível verificar que as ações do *cluster* 1 com 9 casos estão pendentes em diferentes Turmas Recursais.

Figura 7 – Relação das turmas e a quantidade de processos pendentes identificados no *cluster* 1 com 9 casos



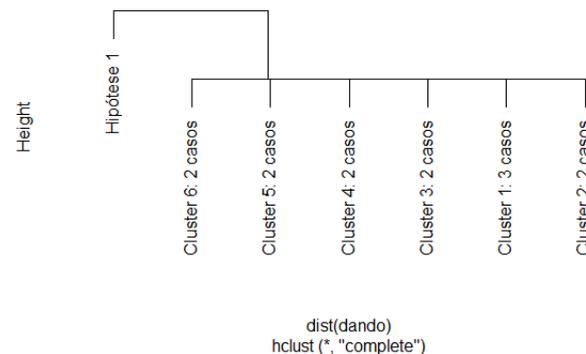
Fonte: os autores (2020).

Esses números indicam que há processos pendentes no Judiciário com petições iniciais iguais distribuídas para diferentes unidades judiciais.

Outra análise foi realizada na 1ª Vara da Fazenda Pública Municipal de Goiânia, sendo encontrados 13 petições iniciais idênticas em números de processos diferentes, distribuídas nos *clusters*, conforme demonstrado no Dendograma da figura 8.

Figura 8 – Representação dos *clusters* encontrados nos casos de similaridade 100% no inteiro teor da petição inicial na 1ª Vara da Fazenda Pública Municipal de Goiânia

**1a Vara da Fazenda Pública Municipal de Gyn**



Fonte: os autores (2020).

### 2.3.2 Casos de 95% de similaridade para o fato e tese jurídica

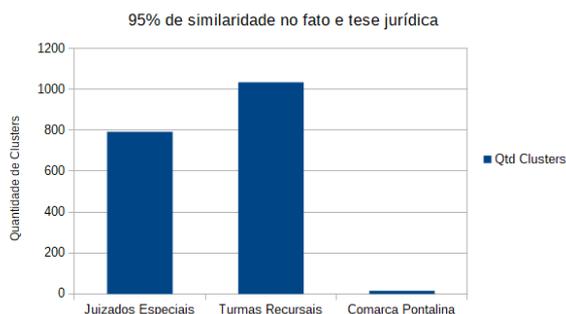
O coeficiente aplicado foi ajustado após estudos de casos realizados, objetivando balancear o maior número de *clusters* encontrados com grau de confiança elevado no cálculo de similaridade dos documentos.

A análise que justificou o uso de 95% no coeficiente, usou duas amostras de petições iniciais: i) 100 (cem) documentos das Turmas Recursais e ii) 100 (cem) documentos dos Juizados Especiais Cíveis de Goiânia. Desse universo de 200 documentos, 50 tinham o mesmo fato e tese jurídica, em processos distribuídos aleatoriamente nessas unidades judiciais. Aplicando o método proposto com coeficiente de 95% de similaridade no fato e tese jurídicos, um *cluster* foi criado com sucesso com 48 documentos dos 50, ou seja, calcula-se neste estudo de caso da amostra uma precisão de 96%. Assim, o coeficiente foi calibrado na Berna com esse valor.

A quantidade de *clusters* gerados pelo método está diretamente ligado ao coeficiente de similaridade utilizado. No método aplicado em produção, foi utilizado 95% de similaridade no fato e tese jurídica com o objetivo de garantir maior exatidão em afirmar que os casos encontrados são iguais.

Na análise da segunda hipótese, os estudos foram realizados em todas as unidades informadas na tabela 2, exceto na 1ª Vara de Fazenda Pública Municipal de Goiânia. Os resultados das análises estão resumidos nas figuras 9 e 10. Na figura 9, são apresentadas a quantidade de *clusters* criados pela Berna em cada unidade judicial.

Figura 9 – Quantidade de *clusters* criados pela Berna



Fonte: os autores (2020).

Outra informação relevante para a análise é saber a quantidade de documentos encontrados, diante do total de processos pendentes. Essa informação é importante para saber se a Berna é ferramenta importante para ajudar no julgamento das ações judiciais.

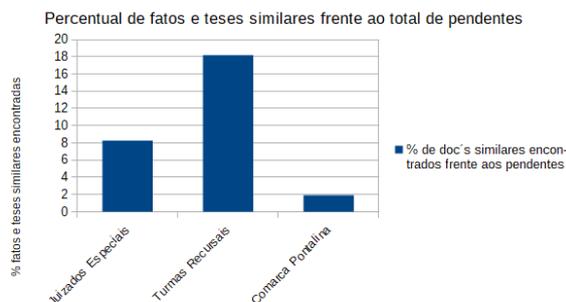
Percebe-se no gráfico da figura 10 que, nas Turmas Recursais, Berna conseguiu identificar situações de conexão em quase 20% dos processos em tramitação. Já nos Juizados Especiais Cíveis da comarca de Goiânia, Berna conseguiu identificar situações de conexão em mais de 8% dos processos em tramitação.

Ainda, identifica-se nos casos encontrados que os processos que guardam essas situações de conexão estão sendo distribuídos para diferentes unidades judiciais nas Turmas Recursais e nos Juizados Especiais Cíveis de Goiânia, podendo, ocasionalmente, o julgador prolatar sentenças diferentes para o mesmos fato e tese jurídicos. Berna então pode avisar as serventias para que elas tomem as providências necessárias.

Outra análise importante é verificar que um pequeno número de *clusters* gerados pela Berna guardam grandes volumes de petições iniciais similares. Pegando apenas os 10 (dez) *clusters* com maior número de documentos similares encontrados, temos que: i) nas Turmas Recursais, eles representam 19% de todos os documentos simi-

lares encontrados, ii) nos Juizados Especiais Cíveis, eles representam 28% de todos os documentos similares encontrados e iii) em Pontalina, eles representam 53% de todos os documentos similares encontrados.

Figura 10 – Apresenta o percentual na relação entre os documentos similares encontrados pela Berna e o total de documentos existentes na unidade judicial



Fonte: os autores (2020).

### 2.3.3 Integração com o Sistema de Processo Eletrônico

Para que a solução tenha efetividade na prática, no dia a dia das unidades judiciais, faz-se necessário implementar solução que permita avisar as serventias dos *clusters* identificados.

O sistema de processo eletrônico ora utilizado pelo Poder Judiciário Goiano é o Projudi, como ferramenta de controle e tramitação de processos. A Berna, após identificação dos *clusters*, gera pendência nos processos que foram distribuídos primeiro em cada serventia que se encontra no *cluster*. Assim, cada secretaria de cartório pode verificar no próprio processo um aviso de pendência. Com o aviso, o serventuário tomará ciência dos casos, podendo tomar as providências necessárias.

A pendência aparecerá tanto na tela principal do secretário do juízo, quanto na tela do próprio processo. Na figura 11 e na figura 12, é possível verificar os locais onde a informação da Berna será visualizada.

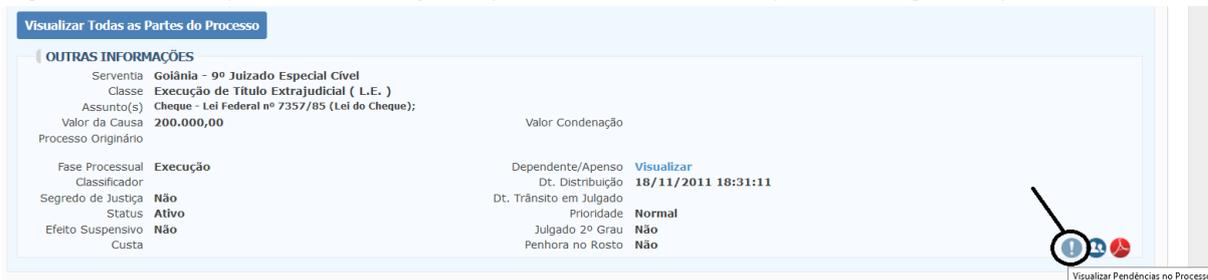
Figura 11 – Tela principal do Projudi, apresentando a pendência gerada pela Berna

PENDÊNCIAS SERVENTIA				
Tipo Pendência	Pré-analisadas Serventia	Não analisadas	Pré-analisadas	Reservadas
Aguardando Prazo Decadencial	0	1	0	0
Alvará	0	1	0	0
Intimação	0	10	0	0
Intimação para Audiência	0	1	0	0
Marcar Audiência	0	35	0	0
Penhora online	0	0	1	0
Verificar Fato e Tese Jurídica	0	2	0	0
Verificar Processo	0	35	0	0
Verificar Petição	0	52	0	0
Verificar Processo	0	3	0	0
Verificar Redistribuição	0	69	0	0
<b>Total</b>	<b>0</b>	<b>242</b>	<b>1</b>	<b>0</b>
Acompanhamento		13		
Expedidas Aguardando Visto		7		
Prazo Decorrido		3		

Fonte: Projudi.

Clicando na pendência, aparece a listagem das ações judiciais similares ao processo em tela, com o número do processo, unidade judicial em que foi distribuído, a data da distribuição e a classe do processo. Exemplo dessa listagem pode ser visto na figura 13.

Figura 12 – Tela do processo no Projudi, apresentando o local da pendência gerada pela Berna



Fonte: Projudi.

Figura 13 – Listagem dos processos no *cluster* relacionados ao processo com a pendência criada, informação gerada pela Berna integrada ao Projudi

Identificados por BERNA (Robô IA332) - Sistema de Inteligência Artificial (Fatos e teses jurídicas 95% similares)						
CASOS PARA TESTE/HOMOLOGAÇÃO - CASOS DE DEMONSTRAÇÃO						
Qnt	Número do Processo	Comarca	Serventia	Classe Processual	Data Distribuição	
1	5450-----	GOIANIA	3º Turma Recursal dos Juizados Especiais	Procedimento do Juizado Especial Cível	xx/xx/xxxx	
2	5451-----	GOIANIA	2º Turma Recursal dos Juizados Especiais	Procedimento do Juizado Especial Cível	xx/xx/xxxx	
3	5451-----	GOIANIA	3º Turma Recursal dos Juizados Especiais	Procedimento do Juizado Especial Cível	xx/xx/xxxx	
4	5451-----	GOIANIA	2º Turma Recursal dos Juizados Especiais	Procedimento do Juizado Especial Cível	xx/xx/xxxx	
5	5451-----	GOIANIA	2º Turma Recursal dos Juizados Especiais	Procedimento do Juizado Especial Cível	xx/xx/xxxx	
6	5451-----	GOIANIA	1º Turma Recursal dos Juizados Especiais	Procedimento do Juizado Especial Cível	xx/xx/xxxx	
7	5451-----	GOIANIA	3º Turma Recursal dos Juizados Especiais	Procedimento do Juizado Especial Cível	xx/xx/xxxx	
8	5451-----	GOIANIA	2º Turma Recursal dos Juizados Especiais	Procedimento do Juizado Especial Cível	xx/xx/xxxx	
9	5451-----	GOIANIA	2º Turma Recursal dos Juizados Especiais	Procedimento do Juizado Especial Cível	xx/xx/xxxx	
10	5451-----	GOIANIA	3º Turma Recursal dos Juizados Especiais	Procedimento do Juizado Especial Cível	xx/xx/xxxx	
11	5451-----	GOIANIA	3º Turma Recursal dos Juizados Especiais	Procedimento do Juizado Especial Cível	xx/xx/xxxx	
12	5451-----	GOIANIA	2º Turma Recursal dos Juizados Especiais	Procedimento do Juizado Especial Cível	xx/xx/xxxx	
13	5451-----	GOIANIA	2º Turma Recursal dos Juizados Especiais	Procedimento do Juizado Especial Cível	xx/xx/xxxx	
14	5458-----	GOIANIA	3º Turma Recursal dos Juizados Especiais	Procedimento do Juizado Especial Cível	xx/xx/xxxx	
15	5459-----	GOIANIA	3º Turma Recursal dos Juizados Especiais	Procedimento do Juizado Especial Cível	xx/xx/xxxx	

Fonte: Berna.

### 3. Conclusão

Na introdução deste artigo, constam as afirmações de que há demandas idênticas em tramitação em unidades judiciais diferentes, algo que tem o condão de macular o princípio do Juiz natural, além disso, nas várias demandas em tramitação, existe número expressivo de processos distribuídos para unidades judiciais distintas que contêm na sua petição inicial o mesmo fato gerador e a mesma tese jurídica, os quais não são identificados pelos atuais mecanismos de detecção de conexão nos sistemas eletrônicos de processos.

A aplicação do método de inteligência artificial proposto permitiu identificar 13 petições iniciais idênticas nas Turmas Recursais, com números de protocolos diferentes na justiça. Três *clusters* foram construídos para reunir esses casos. Os locais de distribuição para o *cluster* de 9 casos foi analisado, sendo verificado que 5 (cinco) processos foram distribuídos para a 4ª Turma Recursal, que 2 (dois) processos foram distribuídos

para a 3ª Turma Recursal e 2 (dois) processos foram distribuídos para a 2ª Turma Recursal. Esses números comprovam a hipótese 1.

Ainda, a aplicação do método proposto possibilitou identificar que 20% dos processos em tramitação nas Turmas Recursais e 8% dos processos em tramitação nos Juizados Especiais Cíveis de Goiânia possuem o mesmo fato gerador e tese jurídica na petição inicial. Vários desses casos foram distribuídos para diferentes unidades judiciais. Esses números comprovam a hipótese 2.

Assim, é possível afirmar que o objetivo deste trabalho foi alcançado, pois o método proposto conseguiu confirmar as hipóteses apresentadas no início deste artigo.

Imagina-se que o modelo de IA e a ferramenta chamada Berna agilizará a movimentação das ações, caso o julgador entenda que possa: i) realizar as conexões devidas nos processos em tramitação; ii) verificar se algum dos *clusters* criados se relacionam aos precedentes, atendendo ao artigo 332

do CPC; iii) analisar a viabilidade de aplicar o Enunciado 73 do FONAJE, nos casos relacionados aos Juizados Especiais; iv) facilitar a identificação dos casos para que as Turmas de Uniformização criem Súmulas e v) estabelecer novas rotinas e gestão dos processos nas unidades judiciais.

Quanto às dificuldades encontradas, tem-se a qualidade dos documentos, petições iniciais, utilizados no processamento do método proposto. A presente ferramenta encontrou problemas na identificação do inteiro teor de algumas peças, não permitindo a extração dos seus caracteres. Percebe-se que alguns documentos são inseridos no ato do protocolo do processo, como imagens, e outros são inseridos sem o cuidado de cadastrar corretamente o seu tipo. Visto que esses documentos são matérias-primas para a Berna, faz-se necessário estabelecer normas mais rígidas no ato de realizar o protocolo dos processos, na parte de inserção das peças iniciais.

Como trabalhos futuros, pretende-se relacionar de forma automática os *clusters* construídos com os casos de precedentes e casos repetitivos já julgados. Ainda, tem-se a intenção de aplicar a Berna para reconhecer e analisar a petição inicial assim que for protocolada, pois poderá já informar, pela pendência, às secretarias dos juízos dos casos de possíveis conexões.

#### 4. Referências

- KIM, R. P. *et al.* **Justiça em Números 2019/ Conselho Nacional de Justiça**. Brasília: CNJ, 2019.
- CASTRO, A. P.; CALIXTO, W. P.; FRANCO, B. F. **Gestão da Informação em Grandes Volumes de Dados no Poder Judiciário**. V Luso-Brazilian Collection – Information Management, Cooperation in networks and Competitiveness, Cap. 2, p. 61-78, 2014.
- SCHAEFER, F. Informática e Organização do Poder Judiciário. **Revista Gestão & Conhecimento**, v.3, n.2, p. 12-24, jul./dez. 2005.
- TEMER, M.; MAGALHAES, A.C. **Modifica o regime e estabelece princípios e regras da Administração Pública**. Casa Civil, Presidência da República, 1998.
- RAMOS, H. S.; ROVER, A. J. **A Lei Administrativa Eletrônica na Perspectiva do Princípio da Eficiência**. II Conferência Sul-Americana de Ciência e Tecnologia aplicada ao Governo Eletrônico-CONEGOV, pp. 33-44, 2005.
- ARSENE, O.; DUMITRACHE, I.; MIHU, I. **Medicine expert system dynamic bayesian network and ontology based**. Expert Systems with Applications 38, 15253-15261, 2011.
- LAMY, J.B. **Owlready: Ontology-oriented programming in python with automatic classification and high level constructs for biomedical ontologies**. Artificial intelligence in medicine 80, 11-28, 2017.
- RANI, M.; DHAR, A. K.; VYAS, O. **Semi-automatic terminology ontology learning based on topic modeling**. Engineering Applications of Artificial Intelligence, 63, 108-125, 2017.
- ZHANG, N. *et al.* **An ontological chinese legal consultation system**. IEEE Access 5 18250-18261, 2017.
- GRUBISIC, A.; STANKOV, S.; PERAIC, I. **Ontology based approach to bayesian student model design**. Expert systems with applications 40, 5363-5371, 2013.
- CECI, M.; GANGEMI, A. **An owl ontology library representing judicial interpretations**. Semantic Web 7, 229-253, 2016.
- FAWEI, B. *et al.* **Using legal ontologies with rules for legal textual entailment, in: AI Approaches to the Complexity of Legal Systems**. Springer, pp. 317-324, 2015.
- CALAMBAS, M. A. *et al.* **Judicial precedents search supported by natural language processing and clustering**. 10th Computing Colombian Conference (10CCC), IEEE, pp. 372-377, 2015.
- ZHANG, N.; WANG, P.; PU, Y. **Challenges and related issues for building chinese legal ontology**. International Conference on Mechatronics, Electronic, Industrial and Control Engineering (MEIC-15), Atlantis Press, pp. 1260-1265, 2015.
- BOUGHANEM, M.; BRINI, A.; DUBOIS, D. **Possibilistic networks for information retrieval**. International Journal of Approximate Reasoning 50, 957- 968, 2009.
- THADA, V; JAGLAN, V. **Comparison of jaccard, dice, cosine similarity coecient to find best fitness value for web retrieved documents using genetic algorithm**. International Journal of Innovations in Engineering and Technology 2, 202-205, 2013.
- JUNIOR, E. S.; SILVA, P. F.; DA SILVA, E. R. G. **Gestão de Conhecimento para Administração Judiciária: Levantamento de demandas de conhecimento e estabelecimento de ontologias**. Revista Democracia Digital e Governo Eletrônico, n. 5, 2011.
- SILVA, E. R. G. **Representation of Judicial**

---

**Knowledge Through Ontologies: Exercise in Electronic Government.** International Conference on Information Systems and Technology *Management*, 2009.

SALTON, G.; MACGILL, M. **Introduction to Modern Information Retrieval.** Computer Science Series, USA. McGraw-Hill, 1983.

BARION, E. C. N.; LAGO, D. **Mineração de Textos.** Revista de Ciências Exatas e Tecnologia, Vol. III, n. 3, 2008.

**Antônio Pires de Castro Júnior**

Graduação em Ciência da Computação pela PUC/GO (1999), mestrado em Ciência da Computação pela UNICAMP (2001), doutorando na Escola de Engenharia Elétrica e Computação da UFG, analista judiciário do TJGO desde 1997.

**Wesley Pacheco Calixto**

Graduação em Física, mestrado em Engenharia Elétrica e de Computação pela UFG, doutorado em Engenharia Elétrica pela UFU, com período na Universidade de Coimbra (UC), pós-doutorado na Carleton University, Ottawa/Canadá.

**Cláudio Henrique Araujo de Castro**

Graduação em Direito pela PUC/GO, especialização em Direito Processual Civil e em Direito Civil pela Faculdade Anhanguera de Ciências Humanas, magistrado titular da 3ª Vara Cível de Goiânia, Juiz Auxiliar da Presidência do TJGO.