



CONSELHO NACIONAL DE JUSTIÇA

Presidente: Ministro José Antonio Dias Toffoli

Corregedor Nacional de Justiça: Ministro Humberto Eustáquio Soares Martins

Conselheiros: Aloysio Corrêa da Veiga

Maria Iracema Martins do Vale

Márcio Schiefler Fontes

Daldice Maria Santana de Almeida

Fernando César Baptista de Mattos

Valtério Ronaldo de Oliveira

Francisco Luciano de Azevedo Frota

Arnaldo Hossepian Salles Lima Junior

Maria Cristiana Simões Amorim Ziouva

André Luiz Guimarães Godinho

Maria Tereza Uille Gomes

Henrique de Almeida Ávila

Secretário-Geral: Carlos Vieira von Adamek

Diretor-Geral: Johaness Eck

Secretaria Especial de Programas, Pesquisas e Gestão Estratégica

Secretário Especial Richard Pae Kim

Juizes Auxiliares Carl Olav Smith

Flávia Moreira Guimarães

Livia Cristina Marques Peres

Revista CNJ / Conselho Nacional de Justiça (CNJ).

- Vol. 1, n. 1 (2015) – Brasília, CNJ, 2015 –

Semestral

Até o v. 2, 2017, publicada anualmente.

e-ISSN 2525-4502

1. Direito – Periódicos I. Brasil, Conselho Nacional de Justiça.

CONSELHO NACIONAL DE JUSTIÇA

SEPN Quadra 514 norte, lote 9, Bloco D, Brasília-DF

CEP: 70760-544

CNPJ: 07.421.906/0001-29

ARTIGOS

Hórus: processamento inteligente dos dados digitalizados da vara de execução fiscal do Distrito Federal

Por Dr. Jairo Simão Santana Melo

Msc. Thiago Arruda Neves

Dr. Weiss Webber Araújo Cavalcante

Resumo: A atividade de digitalização de processos físicos é uma demanda do atual biênio do Tribunal de Justiça do Distrito Federal e dos Territórios (TJDFT), que encerra-se em 2020. A morosidade, falhas e o enfileiramento de digitalizações estão provocando transtornos, principalmente em relação à Vara de Execução Fiscal (VEF), devido ao volume de processos que atualmente são gerenciados por essa unidade judicial. Este artigo apresenta o sistema de processamento inteligente de dados digitalizados da VEF, chamado Hórus, desenvolvido pela equipe de Gestão de Dados do TJDFT (SERGEDE). O Hórus auxilia nas atividades de identificação, classificação, correção, assinatura, carga e registro de novos processos que passarão a tramitar de modo digital. Tudo isso de forma automatizada. Optou-se por estruturar o artigo da seguinte maneira: a) abordagem da problemática que motivou o desenvolvimento do sistema; b) contextualização das iniciativas de inteligência na computação; c) descrição das tarefas desempenhadas pelo Hórus; d) resultados alcançados. A linguagem do artigo, em muitas seções, tem um caráter técnico, no domínio da tecnologia da informação. O sistema desenvolvido, além de incorporar um conjunto de tecnologias de última geração, possibilitou que a atividade jurisdicional fosse realizada com maior eficiência e eficácia, além de auxiliar na arrecadação do governo do Distrito Federal.

PALAVRAS-CHAVE: Inteligência Artificial. Aprendizagem de Máquina. Reconhecimento Óptico de Caracteres. Automação. Processo Judicial Eletrônico.

Abstract: The activity of scanning physical processes is a demand of Tribunal de Justiça do Distrito Federal e dos Territórios' (TJDFT) current biennium, which ends in 2020. The slowness, flaws and queuing of scannings are causing disorders, mainly in relation to Vara de Execução Fiscal (VEF), due to the volume of processes that are currently been managed by this judicial unit. This article presents VEF's intelligent scanned data processing system, called Hórus, developed by the Data Management team of TJDFT (SERGEDE). Hórus assists in the activities of identifying, classifying, correcting, signing, loading and registering of new processes that will start to be processed digitally. All this in an automated way. It was chosen to structure the article as follows: a) approach to the problems that motivated the development of the system; b) contextualization of intelligence initiatives in computing; c) description of the tasks performed by Hórus; d) results achieved. The language of the article, in many sections, has a technical character, in the field of information technology. The system developed, in addition to incorporating a set of state-of-the-art technologies, made it possible for the jurisdictional activity to be carried out with greater efficiency and effectiveness, besides helping in the collection of Distrito Federal's government.

KEYWORDS: Artificial Intelligence. Machine learning. Optical Character Recognition. Automation. Electronic Judicial Process.

1 INTRODUÇÃO

A Lei nº 11.697, de 13 de junho de 2008, que dispôs sobre a organização da Justiça do Distrito Federal e dos Territórios, em seu artigo 35, previu que compete ao Juízo da Vara de Execução Fiscal "processar e julgar as execuções em que o Distrito Federal ou entidades de sua administração descentralizada, inclusive empresas públicas e sociedades de economia mista de que participe, forem autores, réus, assistentes, litisconsortes, intervenientes ou oponentes, excetuadas as de falência, acidentes de trabalho e de meio ambiente, desenvolvimento urbano e fundiário".

Nesse diapasão, o pleno administrativo do Tribunal de Justiça do Distrito Federal e dos Territórios - TJDFT editou a Resolução nº 19, de 2 de dezembro de 2009, dispondo sobre a instalação da Vara de Execução Fiscal, com competência territorial abrangendo todo o Distrito Federal e competência material/funcional para "processar e julgar as ações de execução de dívida ativa do Distrito Federal e de suas autarquias, bem como os embargos à execução a

elas correspondentes", conforme se extrai de seus artigos 1º, parágrafo único, e 2º.

Em suma, a Vara de Execução Fiscal do Distrito Federal - VEF tem a função de processar e julgar os executivos fiscais relativos à dívida ativa da Fazenda Pública do Distrito Federal, compreendendo a tributária e a não tributária, abrangendo a atualização monetária, juros e multa de mora, bem como demais encargos previstos em lei ou contrato, segundo preconizado pelo artigo 2º, §§ 1º e 2º, da Lei nº 6.830, de 22 de setembro de 1980 (Lei de Execução Fiscal).

O atual perfil da VEF pode ser analisado a partir de dados levantados pelo Conselho Nacional de Justiça - CNJ, a fim de contextualizar a posição daquela unidade jurisdicional frente a outras de mesma natureza, considerando-se ainda seu atual acervo processual, bem como seus recursos humanos e materiais disponíveis.

O CNJ publicou o Relatório Justiça em Números 2018

(ano-base 2017), sendo destacado o impacto negativo gerado pela fase de execução nos dados de litigiosidade do Poder Judiciário brasileiro, considerando um acervo do Poder Judiciário brasileiro de 80,1 milhões de processos pendentes de baixa no final de 2017 em que 53% se referiam à fase de execução. Apontou, ainda, que, deste estoque, 74% são compostos por execuções fiscais, responsáveis pela alta taxa de congestionamento do Poder Judiciário, representando aproximadamente 39% do total de casos pendentes.

No caso do TJDFT também não é diferente tal constatação, pois a VEF já nasceu com um número extremamente elevado de processos, cerca de 186.000 (cento e oitenta e seis mil), em face da redistribuição de processos das 8 (oito) varas de fazenda pública, dos quais, aproximadamente, 90.000 (noventa mil) ainda não tinham sido autuados.

Em fevereiro de 2019 a VEF detinha um acervo processual de mais 300.000 processos, pois havia 254.431 processos em tramitação (163.758 físicos e 90.673 eletrônicos), segundo informações do Boletim Estatístico Mensal - 1ª Instância, e mais de 90.000 processos (físicos e eletrônicos) arquivados provisoriamente em face da incidência do Provimento 13/2012 da Corregedoria do TJDFT.

Ademais, dentre os processos digitalizados, há os que foram distribuídos e os que ainda restam pendentes de distribuição, tendo a Secretaria de Distribuição e Movimentação de Autos Judiciais da 1ª Instância do TJDFT informado que representam 39.812 processos.

Vale destacar que, segundo o mencionado Relatório Justiça em Números, a taxa nacional de congestionamento nas execuções fiscais em 2017 foi de 91,7%, sendo na Justiça Estadual de 91%. No caso do TJDFT, a taxa de congestionamento chega a 95%, a maior dentre os tribunais de médio porte.

Para o enfrentamento do acervo processual mencionado, a VEF dispõe, atualmente, de apenas um juiz titular, 3 juízes substitutos, 35 servidores e 43 estagiários.

Diante desse quadro, não distinto do apresentado por relatórios anteriores do CNJ, surgiu a necessidade de uma atuação diferenciada na VEF, pautada por uma gestão judiciária estratégica, que tenha em vista as relações institucionais da Vara com a Administração Superior do TJDFT e órgãos externos (como a Procuradoria do Distrito Federal), uma gestão administrativa (recursos humanos e processos de trabalho) e uma atuação jurisdicional que busque uniformidade e segurança jurídica, o que tem sido implementado desde o final de junho de 2017, com a remoção para a VEF do novo juiz titular.

A utilização de mecanismos que objetivem a efetividade do processo de execução fiscal e sua movimentação, tais como a classificação de processos em massa e repetitivos; a realização de conciliações fiscais, a priorização no tratamento de atos constritivos, sentenças extintivas e análise de embargos à execução; a adoção de decisões concentradas; a análise de casos semelhantes em sede de exceção de pré-executividade; e a desjudicialização dos procedimentos de cobrança têm se revelado medidas úteis.

O fato é que, ainda assim, os recursos materiais convencionais demonstraram ser obsoletos, o que tornou

necessária a busca por soluções tecnológicas, uma vez que são as ferramentas mais adequadas quando há a necessidade de se praticar atos repetitivos e de menor complexidade, como nos casos de realização de citação, de decisões que determinam constrições por meio de sistemas eletrônicos, de decisões suspensivas do processo em razão do parcelamento administrativo, de sentenças extintivas em razão do pagamento ou cancelamento do crédito fiscal, dentre outros atos.

Exemplo disso foi a viabilização da participação da VEF na Semana Nacional de Conciliação, nos dias 20, 21 e 22 de novembro de 2017, por meio de Acordo de Cooperação Técnica, firmado entre o TJDFT e o Distrito Federal, que possibilitou a expedição de mais de 6.000 mandados de citação em tempo recorde em razão da automação da expedição.

Outras situações que mereceram a utilização da automação, desenvolvidas no âmbito do próprio Tribunal, foram o registro do teor do requerimento de desistência e das sentenças, baixa dos processos, realização de *checklist* e posterior arquivamento no caso das execuções fiscais extintas em virtude da incidência do Decreto 38.650/2017 editado pelo Governador do Distrito Federal, resultando no tratamento de mais de 35.000 processos em apenas um mês.

Destaca-se, ainda, o lançamento de andamentos nos processos remetidos à digitalização terem sido realizados de forma automatizada ante o fornecimento de listagem de processos pela VEF, permitindo a alocação dos recursos humanos para outras tarefas.

Nessa esteira, vislumbram-se diversos outros atos processuais que poderiam ser realizados por meio de ferramentas tecnológicas, inclusive despachos e decisões de menor complexidade que, atualmente, exigem tempo e energia de servidores que poderiam ser melhor canalizados para tarefas mais complexas.

É o caso, por exemplo, das decisões que determinam bloqueio/penhora/arresto e necessitam serem efetivadas por meio do BacenJud. Referida tarefa é realizada, de forma simplista, a partir de variáveis objetivas (citação válida, não pagamento, requerimento de bloqueio) e cujo desdobramento importa no registro da decisão e alimentação de dados no BacenJud para efetivação do bloqueio. Cerca de 5 servidores são os responsáveis por minutar tais decisões, com uma produtividade diária e individual de 30 processos, enquanto uma máquina poderia fazê-lo em maior quantidade, num ínfimo período de tempo e com acurácia próxima de 100%.

Mencionem-se, ainda, outros atos repetitivos, também solucionáveis com tecnologias adequadas, tais como a expedição de mandados de citação, decisões suspensivas do processo em razão do parcelamento administrativo, sentenças extintivas em razão do pagamento ou cancelamento do crédito fiscal, decisões recorrentes em sede de exceções de pré-executividade (ex.: analisando a prescrição, cujas balizas são objetivas) e demais atos cartorários (certidões, verificação no Sistema de Tributação e Administração Fiscal da Secretaria da Fazenda do DF – SITAF, verificação de enquadramento no Provimento 13/2012 ou no Decreto 38.650/2017, dentre outros).

Para tanto, além das ferramentas de automação, a Inteligência Artificial – IA seria capaz de aplicar regras lógicas a um conjunto de dados, aprender com os erros e acertos, reconhecer padrões visuais/sensoriais/comportamentais e aplicar o raciocínio aos casos concretos, em princípio repetitivos e de menor complexidade. Frise-se que ferramentas baseadas na IA em nada retirariam o poder de autonomia do magistrado e dos servidores, uma vez que sua calibragem sempre estaria em consonância com o entendimento e supervisão daqueles.

É certo que a adoção de Inteligência Artificial no processo judicial tributário, além de subsidiar a gestão judicial estratégica de processos em massa e repetitivos de execução fiscal, daria efetividade ao princípio constitucional da eficiência na administração pública, pois otimizaria sobremaneira os recursos humanos e materiais colocados à disposição.

Possibilitaria, ainda, que o TJDFT cumprisse a sua missão de “proporcionar à sociedade do Distrito Federal e dos Territórios o acesso à justiça e a resolução dos conflitos, por meio de um atendimento de qualidade” e alcançasse o objetivo de “até 2020, apresentar resultados que reflitam o aumento da produção, da eficiência e da qualidade”, sem descuidar dos valores da “imparcialidade, ética, celeridade, credibilidade e transparência”.

Considerando o acervo processual da VEF representar cerca de 1/3 (um terço) do acervo do TJDFT, a tecnologia daria ensejo ao alcance das metas propostas pelo CNJ em menor tempo, reduzindo custos e melhorando o atendimento ao jurisdicionado.

Outrossim, propiciaria a recuperação de parte do crédito fiscal da Fazenda Pública do Distrito que, segundo informações da Procuradoria-Geral do Distrito Federal, totaliza 32 bilhões de reais, que já são objeto de cobrança por meio dos executivos fiscais ajuizados na VEF.

Este artigo tem como objetivo relatar a experiência em projeto embrionário de Inteligência Artificial no TJDFT com o intuito de acelerar a virtualização do acervo de processos da VEF, bem como apresentar o cenário tecnológico atual e os materiais e métodos utilizados para alcançar os objetivos pretendidos.

O artigo que segue está estruturado em quatro seções. A primeira seção contextualiza o cenário atual de tecnologia

da informação, principalmente com relação a iniciativas de inteligência artificial. A segunda seção aborda as tarefas desempenhadas pelo Hórus, destacando conceitos, materiais, métodos e tecnologias relacionadas. A terceira apresenta os primeiros resultados alcançados com a execução do sistema na VEF. Por fim, na conclusão, são resumidos os principais destaques envolvidos na realização do projeto do sistema Hórus.

2 CENÁRIO ATUAL DE TECNOLOGIA DA INFORMAÇÃO

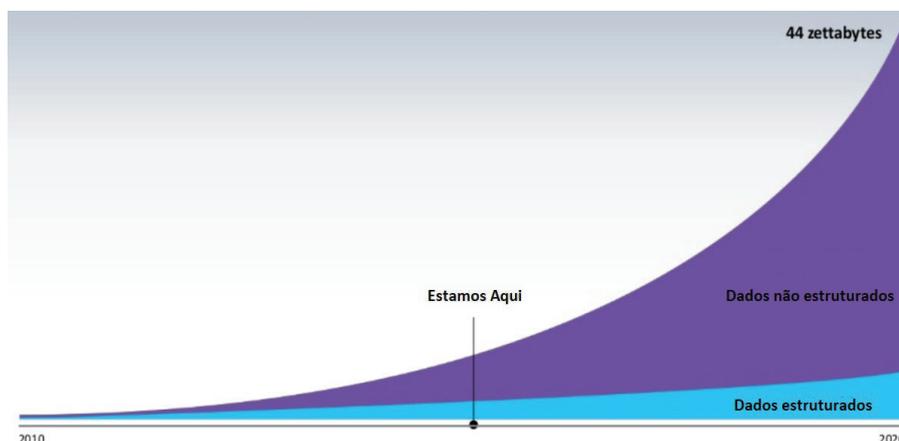
Nos últimos anos, a corrida pela transformação digital tornou-se acirrada em todos os nichos de mercado. As corporações perceberam que, ao aliar a tecnologia da informação e o negócio, teriam maior qualidade, eficiência e lucro na prestação de serviços. A Internet passou de uma mera rede mundial de computadores para uma verdadeira plataforma de prestação de serviços digitais.

Hoje, bancos virtuais, lojas on-line, redes sociais, são aplicações essenciais no cotidiano das pessoas. Conforme Vaz (2009, p. 63) previu há uma década atrás: “não saber usar a internet em um futuro próximo será como não saber abrir um livro ou acender um fogão, não sabermos algo que nos permita viver a cidadania na sua completude”.

Em meio a esta revolução digital, os negócios tiveram que se reinventar em um contexto no qual a folha de papel se tornou a grande vilã da produtividade. Agora o acesso aos serviços tornou-se mais fácil e útil, sem burocracias, o que naturalmente resultou em maior procura pela sociedade. Nos órgãos, o caminho não foi diferente devido à necessidade de redução dos gastos e aumento da celeridade na prestação de serviços públicos. Nesse sentido, várias instituições priorizaram a digitalização dos processos e a oferta de serviços digitais, extinguindo balcões de atendimento, filas, galpões de arquivamento e impressoras. Foi o caso, por exemplo, do Superior Tribunal de Justiça que, em meados de 2009, alavancou um grande projeto de digitalização de todo acervo de processos, influenciando vários outros tribunais.

O efeito dessa tendência foi um aumento exponencial do volume de informações digitais dentro das organizações - muitas vezes não estruturadas - chegando ao ponto de não serem mais suportadas pelos sistemas transacionais clássicos, como pode ser observado na Figura 1.

Figura 1 - Projeção de volume de dados não estruturados até 2020



Fonte: HECKSTER, 2015

Novas soluções tecnológicas surgiram no sentido de classificar, processar e analisar os dados massivos (*Big data*) de forma inteligente, a fim de produzir informações úteis aos gestores e usuários. Dentre elas, destacam-se os softwares de mineração e análise de dados, as plataformas de *Business Intelligence* e as técnicas de Inteligência Artificial (IA). Cabe destacar as ferramentas de *Data Warehouse*, que auxiliam na consolidação dos vários silos de dados em um ambiente centralizado e organizado, propício para a análise das informações.

A Inteligência Artificial já é uma realidade em nossas atividades cotidianas. Consultas inteligentes realizadas em sites de buscas, onde a informação procurada tem relação com o contexto em foco, compras direcionadas e análise de perfil com recomendações baseadas nas interações e preferências do cliente, processamento de voz humana, atendimento online por softwares especialistas - os *chatbots* - e assistentes de rotas por geolocalização estão presentes em vários serviços digitais.

Apesar de a Inteligência Artificial ter ganhado bastante notoriedade mais recentemente com um grande número de aplicações comerciais, ela não é um campo de estudo novo. Suas raízes remontam do ano de 1950, quando o cientista Alan Turing realizou os primeiros ensaios que avaliaram a capacidade de pensar de uma máquina (TURING, 1950). A IA tornou-se um ramo de pesquisa sólido da ciência da computação que busca, através de algoritmos e processos estatísticos, construir mecanismos e/ou dispositivos que simulem a capacidade do ser humano de aprender, pensar e resolver problemas, ou seja, de ser inteligente.

Segundo Shannon, Mayer e Adelsberger (1985, p. 275), Inteligência Artificial "é o estudo de como os computadores podem realizar tarefas que hoje são melhor desempenhadas pelas pessoas". De forma mais pragmática, Waterman (1986, p. 419) define a IA como "uma subárea da Ciência da Computação que objetiva desenvolver softwares computacionais inteligentes. Esses softwares são solucionadores de problemas, melhoram sua própria performance, interpretam linguagens, reconhecem esquemas visuais e, enfim, comportam-se de maneira que seria considerada inteligente se observada num homem".

- A IA pode atuar em dois eixos maiores: a imitação das habilidades humanas - gestos, visualização, comunicação - ou a replicação dos resultados estabelecidos pelo homem através de sua habilidade e experiência (SHANNON; MAYER; ADELSBERGER, 1985). Rich et al (1994, p. 72) apresentam um levantamento dos principais campos de aplicação da IA:
- Processamento de Linguagem Natural: objetiva a construção de softwares capazes de compreender a linguagem natural e gerar textos, com o intuito de aprimorar a comunicação entre o homem e o computador. Dentre as aplicações, encontramos atualmente os *chatbots* e tradutores como o *Google Translate*;
- Reconhecimento de Padrões: permite que os softwares sejam capazes de reconhecer a fala em linguagem natural, - os caracteres digitados ou manuscritos - reconhecer objetos e imagens, dentre outros;
- Visão Computacional: é o processo de replicação

da visão humana usando software e *hardware*. A visão computacional é uma disciplina que estuda como reconstruir, interromper e compreender uma cena 3D a partir de suas imagens 2D em termos das propriedades da estrutura presente na cena;

- Aprendizado de Máquina (*Machine learning*): dá aos computadores a habilidade de aprender sem terem sido explicitamente programados com esse conhecimento;
- Robótica e Percepção: criação de equipamentos e softwares que emulam os comportamentos humanos como, por exemplo, os veículos autônomos que possuem computadores de bordo com habilidades para dirigir, acelerar e frear .

Cabe aprofundar duas dessas aplicações de IA que foram amplamente utilizadas no Projeto Hórus - o Reconhecimento de Padrões e o Aprendizado de Máquina. Para converter as imagens de documentos digitalizados para um formato que possa ser interpretado e processado pelo computador, foi utilizada a tecnologia de OCR (*Optical Character Recognition*), ou reconhecimento óptico de caracteres. O OCR permite identificar os caracteres por meio de um arquivo de imagem contendo textos escaneados ou manuscritos, empregando técnicas como redes neurais artificiais, classificação por aprendizagem supervisionada, dentre outras, que permitem maximizar o reconhecimento, mesmo em situações com diferentes fontes e idiomas, distorções e falta de qualidade na imagem. Nesse sentido, foi utilizada uma biblioteca em software livre denominada *Tesseract*, que atualmente é mantida pelo *Google*, sendo um padrão de fato no processamento de OCR de documentos digitais (SMITH, 2007).

Por meio da IA, um programa de computador adquire a capacidade de melhorar sua performance em determinada tarefa com base na experiência em executá-la. Isso é conhecido como Aprendizado de Máquina. Existem duas divisões principais da forma pela qual o aprendizado de máquina pode ocorrer: supervisionado e não supervisionado.

Pode ser usado o aprendizado supervisionado, - quando um conjunto de dados de treinamento já classificado é apresentado e cabe ao algoritmo aprender os padrões dessas classes para que avalie corretamente novos conjuntos de dados - ou o aprendizado não supervisionado - quando não se tem ideia das classes (ou rótulos) existentes no conjunto de dados e o algoritmo faz agrupamentos por similaridade dos registros, possibilitando que a aplicação determine as classes. O aprendizado supervisionado trabalha com dados rotulados, ou seja, dados nos quais a saída (resultado) desejada é conhecida. Essa saída possui relação com as características dos dados. Nesse tipo de aprendizado, há um "professor" que avalia a resposta da máquina de acordo com os dados de entrada fornecidos. Portanto, o objetivo é aprender a regra geral que relaciona entrada e saída. Problemas de aprendizado supervisionado são divididos, em sua maioria, em Classificação e Regressão.

Por outro lado, o aprendizado não supervisionado trabalha com dados que não possuem um rótulo, ou seja, não se sabe qual é a saída desejada. Nesse caso, a máquina deve aprender sozinha as relações entre entrada e saída, ou seja, descobrir padrões, regularidades, categorias/perfis.

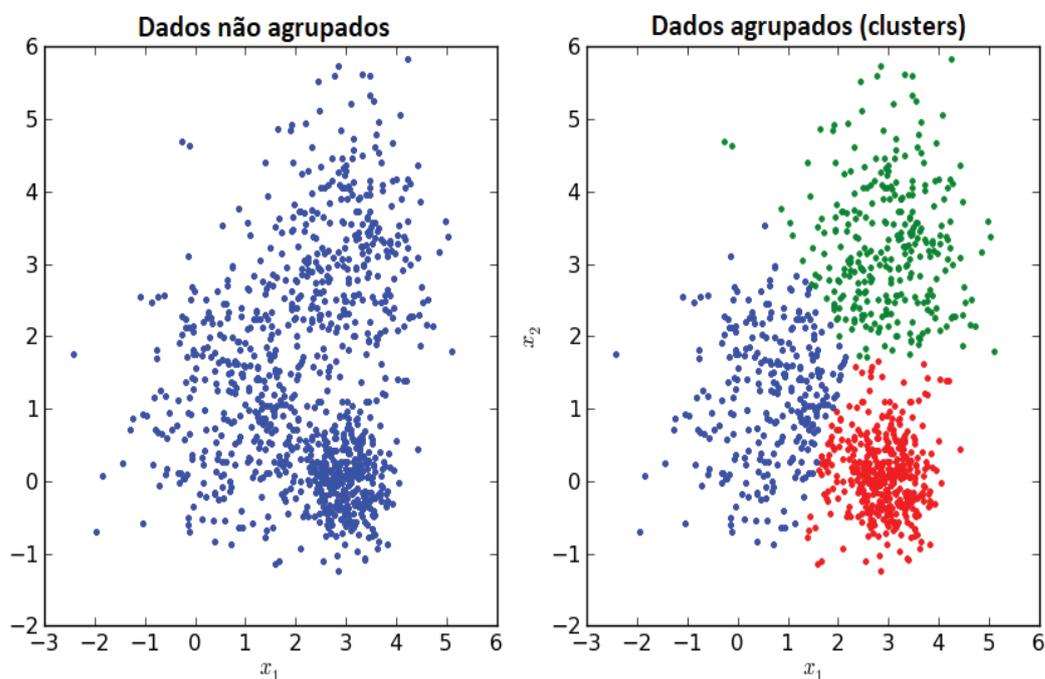
Problemas de aprendizado não supervisionado incluem, em sua maioria, abordagens de agrupamento (*clustering*).

A Inteligência Artificial já é uma realidade também no Poder Judiciário (projetos no STF, TJPE, TJRO). A maioria das iniciativas da Justiça está voltada para a classificação de modo supervisionado, isto é, existe a necessidade de que um especialista gerencie os atributos do processamento para garantir a efetividade do mesmo. Ou seja, alguém deve fornecer os rótulos dos dados (as saídas desejadas) para treinamento da máquina. Por exemplo, ao trabalharmos com um conjunto de dados de documentos processuais, um magistrado pode informar quais características do documento (formato, cabeçalho, palavras-chave) indicam se este é uma Certidão ou um Ofício.

Um caso de sucesso nesse sentido é o projeto Victor, do STF, utilizado inicialmente para processar os recursos extraordinários que sobem para o STF e para identificar quais estão vinculados a determinados temas de repercussão geral.

Por exemplo, no caso do projeto Hórus, para classificar automaticamente os documentos, foi utilizado o algoritmo de aprendizagem *K-Means*. Essa tecnologia tem o objetivo de separar grupos de dados (*clusters*) por meio da avaliação da similaridade dos atributos dos registros, de forma não supervisionada, ou seja, não é necessário treinamento prévio com registros já rotulados. Isso é realizado por meio do cálculo da distância média de cada registro em relação aos núcleos de maior densidade, denominados centróides. A Figura 2 apresenta um exemplo de agrupamento de dados após a aplicação do *K-Means*.

Figura 2 - Exemplo de agrupamento de dados por meio do *K-Means*.

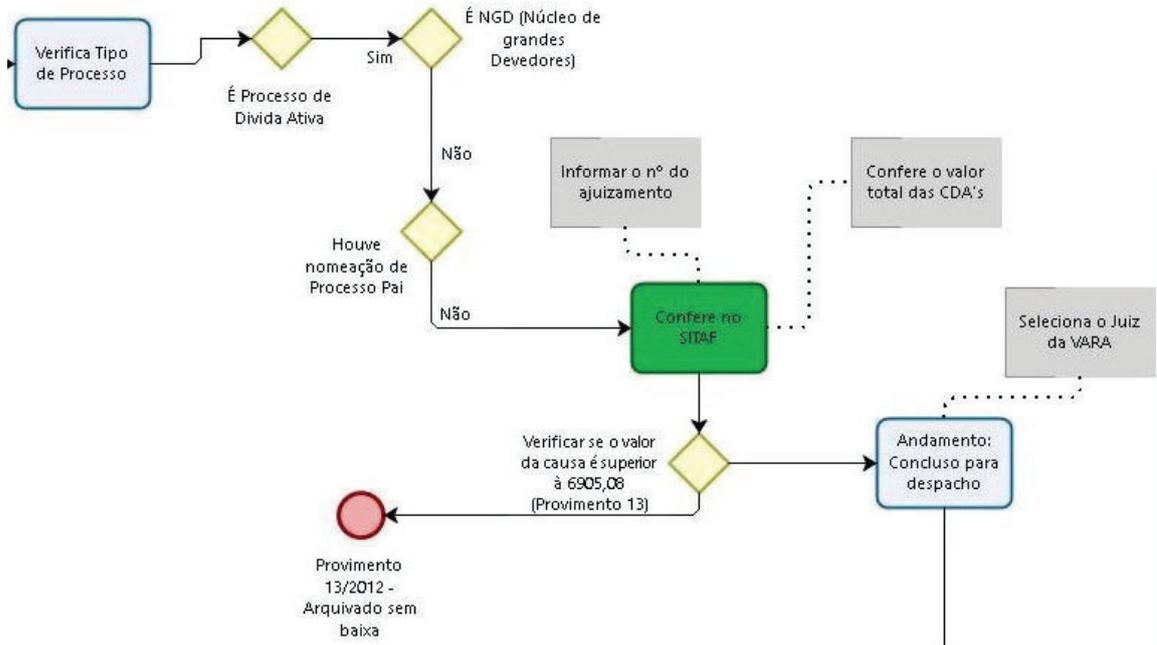


Fonte: VISWARUPAN, 2019

3 TAREFAS DO HÓRUS

O processo judicial eletrônico já é uma realidade no TJDF. Atualmente, cerca de 10% dos processos do tribunal já tramitam nesta plataforma. Basicamente, esse sistema digital operacionaliza o fluxo de trabalho da unidade, como por exemplo o fluxo de atividades exercidas pela secretaria VEF, representado pelo diagrama de processo de negócio descrito na Figura 3.

Figura 3 - Notação BPMN do fluxo de trabalho da da secretaria da VEF



Atualmente, a VEF representa 1/3 dos processos do TJDF. Nesse sentido, carece de uma automatização da totalidade de suas atividades, quando possível, a fim de que a celeridade processual seja uma característica dessa unidade judicial. Entretanto, essa unidade ainda dispõe em seu acervo de processos físicos que tramitam de modo 100% manual, dificultando a celeridade e consumindo recursos, como a impressão dos autos do processo.

Uma das diretrizes do atual biênio do TJDF (2018-2020) é justamente a digitalização de processos. Contudo, as atividades posteriores a essa fase de digitalização são

manuais e promovem um afinamento na fase de cadastro e distribuição de processos no sistema eletrônico, pois este depende da atuação de servidores na importação dos diferentes tipos de documentos e meta-informações do processo.

Nesse sentido, vislumbrou-se uma solução automatizada e inteligente, denominada Hórus, que auxilia nas atividades de identificação, classificação, correção, assinatura, carga e registro dos novos processos que passarão a tramitar de modo digital. O fluxo de processamento pode ser observado na Figura 4.

Figura 4 - Fluxo de tarefas automatizadas e inteligentes do Hórus



3.1 DIGITALIZAÇÃO E OCR

A primeira atividade do fluxo é a digitalização dos volumes físicos do processo, a qual configura-se como uma etapa rápida, caso seja utilizado *hardware* especializado para essa tarefa. Atualmente, o TJDF já dispõe de dois centros de digitalização, incumbidos de atender a meta

estabelecida pela administração. Uma subtarefa dessa fase consiste na fragmentação e indexação dos diferentes tipos de documentos que fazem parte dos autos do processo. No entanto, por ser uma atividade manual, algumas falhas na classificação foram identificadas.

Após a digitalização, foi identificada a necessidade de

reconhecimento óptico de caracteres, possibilitando a localização de certas características dos processos, como número de identificação, informações das partes e informações básicas do processo de modo automatizado. Para isso, foram utilizadas as melhores práticas, como o uso da biblioteca *Tesseract*. Nesse sentido, algumas simulações foram realizadas nas peças processuais fragmentadas e indexadas da primeira fase, como pode ser observado nas Figuras 5 e 6, a fim de verificar a viabilidade do uso da biblioteca.

A efetividade dessa segunda fase nos processos da VEF foi alcançada com êxito, devido à padronização dos documentos que configuram a petição inicial do processo, a qual pode ser representada por uma certidão de digitalização ou pela Certidão de Dívida Ativa (CDA). Após o OCR do documento que identifica a certidão, é possível observar a presença do número que identifica o processo nos sistemas legados do TJDF.

Figura 5 - OCR do documento que identifica a certidão de digitalização

```
TRIBUNAL DE JUSTIQA DO DISTRITO
FEDERAL E DOS TERRITORIOS (INTERNO)

Registrado sob o N9 Unico 2013011-16.3655.6..
(20130111636556)

CERTIDAO DE VALIDAVO

Certifico que os autos eletrônicos correspondem aos
fl'sicos, adquirindo suas páginas nova numeragéo eletrônica.

Brasilia, 14 de maio de 2018.

TRIBUNAL DE JUSTIQA DO DISTRITO FEDERAL E DOS
TERRITORIOS (INTERNO)

(*) Documento assinado eletronicamente
por (00259)/Tatiane Rodriques de Azevedo nos termos
do Ar1.19 §29 inciso ||| alinea "b" da Lei 11.419/2006
```

3.2 CONSULTA DE METADADOS

O passo seguinte, após o reconhecimento dos caracteres dos documentos digitalizados, consiste na identificação de algumas informações relevantes para as etapas seguintes do fluxo. Primeiramente, o número de identificação do processo é reconhecido pela certidão de digitalização, como pode ser observado na Figura 4. Caso esse número não seja encontrado, o respectivo nome do diretório, o qual armazena as peças processuais que foram fragmentadas, corrobora com essa identificação. Outras informações também podem ser recuperadas e processadas a partir da CDA, descrita na Figura 6, como por exemplo: a) informações do pólo passivo, tais como nome, CEP, endereço; b) informações que identificam a CDA; c) informações que identificam a dívida atualizada na data de expedição do documento.

Essas informações podem contribuir para o passo

posterior, que consiste em recuperar os metadados do processo a partir do *Data Warehouse* corporativo denominado DWJURIS, através de um serviço que retorna as informações processuais baseadas no padrão CNJ de formatação de processos, permitindo a interoperabilidade entre sistemas, como pode ser observado na Figura 7. Esse padrão procura identificar as informações básicas de processos, tais como: classe, assunto, circunscrição, prioridades, pólos processuais, magistrados, advogados, órgão julgador, movimentos e documentos.

O *Data Warehouse* corporativo do TJDF está em construção há 3 anos e várias técnicas de melhoria de dados foram utilizadas para que os sistemas demandantes pudessem usufruir de dados já tratados. Dentre essas técnicas e correções, é possível destacar o tratamento de dados em relação a advogados e incidência penal, que constantemente eram alvo de questionamentos por parte de relatórios do CNJ.

Figura 6 - OCR do documento que identifica a CDA

```

NUMERO DATA ORDEM NATUREZA E
1 0162082690 02/10/13 DE001250008772012 0136-0136/2009-0100 8
2 0162082703 02/10/13 DE001250008772012 0136-0136/2009-0100 -n
3 0162082711 02/10/13 DE001250008772012 0136-0136/2009-0100
4 0162082720 02/10/13 DE001250008772012 0136-0136/2009-0100
5 0162082738 02/10/13 DE001250008772012 0136-0136/2009-0100
6 0162082746 02/10/13 DE001250008772012 0136-0136/2009-0100 i
7 0162082754 02/10/13 DE001250008772012 0136-0136/2009-0100
E .
(f1 DATA CONSTITUIÇÃO 0 D VALOR ORIGINAL EM R$ K M . . .
DEFINMVA PRINCIPAL MULTA CORRE( ésgglfFTARIA JUROS DE MORA (LC 435/01)
1 22/07/09 1 166,99 16,89 0,00 0,00
2 28/08/09 2 261,87 26,18 0,00 0,00
3 30/09/09 3 149,76 14,97 0,00 0,00
4 04/11/09 4 235,71 23,57 0,00 0,00
5 03/12/09 5 207,69 20,76 0,00 0,00
6 V 11/12/09 6 191,44 19,14 0,00 0,00
7 19/01/10 7 169,23 16,92 0,00 0,00
8 j @ / 2
2r .
- , 1 3 4 5 6 7
CORREÇÃO MONETARIA 20/07/09 20/08/09 21/09/09 20/10/09 20/11/09 21/12/09 20/01/10
JUROS DE MOR A 1 20/07/09 2 20/08/09 3 21/09/09 4 20/10/09 5 20/11/09 6 21/12/09 7 20/01/10 8
k ' J
EL ~ G CERTIFICO QUE 0(8) DEBITO(S) DISCRIMINADO(S) NESTA CERTIDAO CONTEM(EM) TODOS
CERTIDAO N° 08 ELEMENTOS DO(S) SEU(S) TERMO(S) DE INSCRIVO NO REGISTRO DE DIVIDA ATIVA DO
DISTRITO FEDERAL
0005491886 . . . . . 9%.. $9229
\ Em: 30/10/2013 Chefe do Nucleo de Gestáo da Divida Ativa da PGDF

Requer, assim, se digne ordenar a citação, segundo as diversas modalidades legais, da(s) pessoa(s)
pelas sucessivas formas previstas no art. 8º, da Lei nº6.830/80, para que pague(m) incontinenti
e juros (LC 435/01) e demais acessórios de direito, ou garanta(m) a execução, para efeitos de
de não serem encontrados os devedores para citação por oficial de justiça, nos endereços acima
de ausência de pagamento ou garantia da execução, requer seja determinada a imediata penhora d
preferencialmente, a penhora de dinheiro em espécie, em depósito ou aplicação em instituição l

I, da Lei nº 6.830/80). Se a diligência do Bacenjud for negativa, requer vista dos autos para
Brasília ' DF de de 2013 Da-se a presente valor fiscal de

R$: 1 .523,12

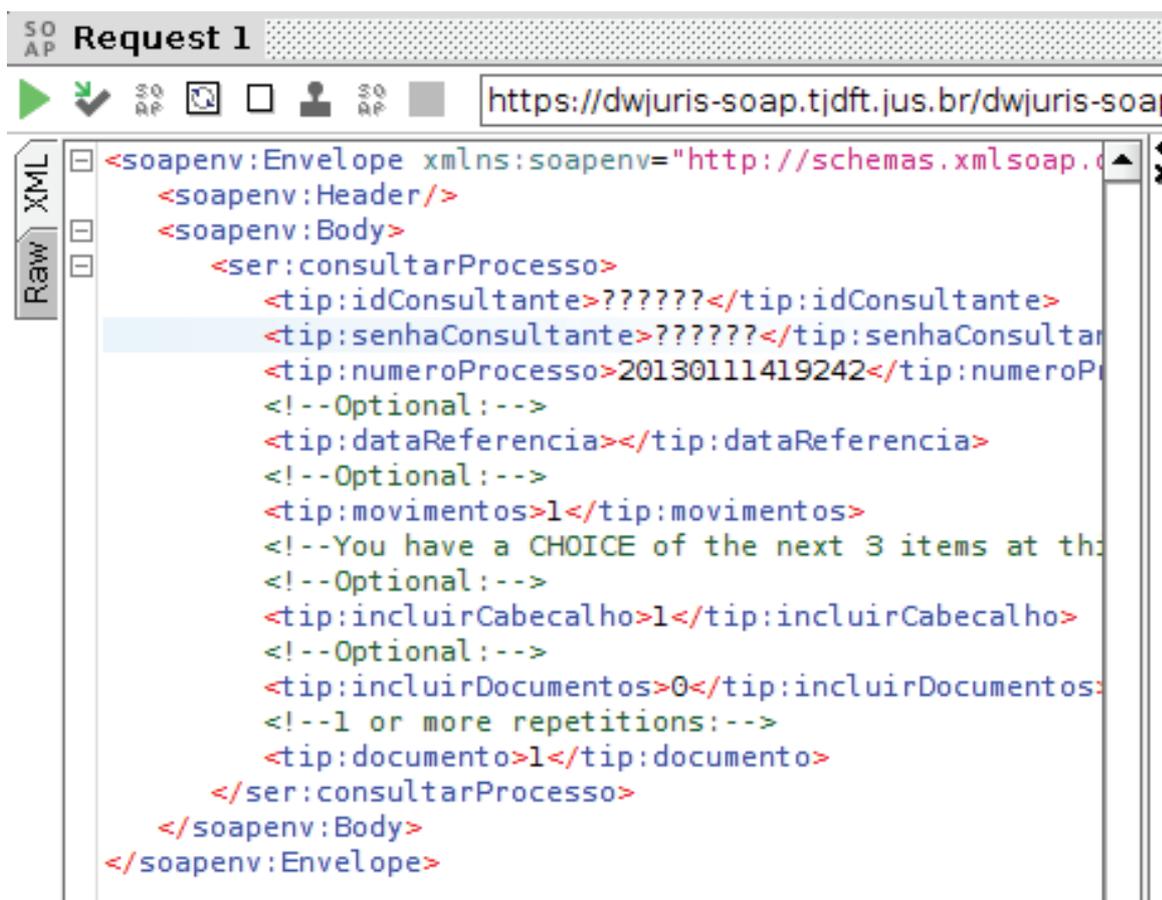
Valor atualizado nesta data.

BRUNO PAIVA DA FONSECA R$2 3.013,56

```

Atualmente, o DWJURIS é parte integrante de diversos sistemas do TJDF, devido ao tratamento, padronização e unificação dos dados processuais em uma única base, permitindo o acesso de modo centralizado. Dessa forma, o projeto Hórus também usufrui dessas melhorias, porém algumas informações necessitam ser equalizadas com base no padrão esperado pelo PJE, como: código do órgão julgador e código de localidade do processo, possibilitando o respectivo cadastro de forma automática.

Figura 7 - Interface de Interoperabilidade disponibilizada pelo DWJURIS para consulta de processos baseada no padrão MNI do CNJ



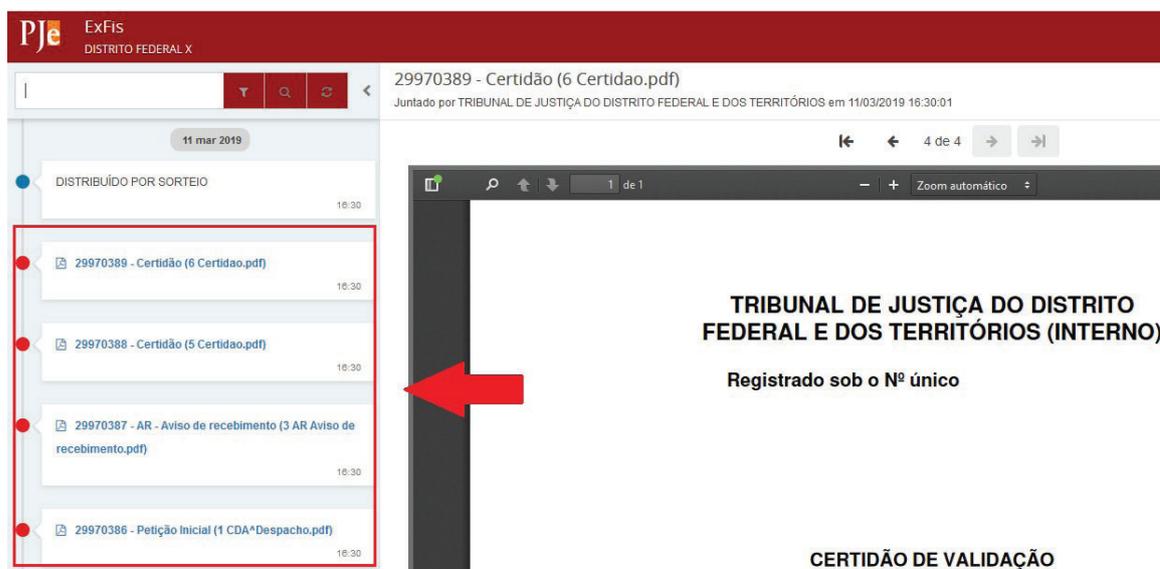
3.3 CLASSIFICAÇÃO

A etapa posterior à recuperação dos metadados do processo, identificada como depara/correção e classificação, objetiva preparar um conjunto de dados para ser inserido no PJE. A subatividade de classificação faz uso de algumas técnicas de IA, as quais serão detalhadas nas seções seguintes.

3.3.1 APRENDIZADO DE MÁQUINA APLICADO AO HÓRUS

Para a equipe do TJDF, o desafio posto na etapa de classificação do fluxo ilustrado na Figura 3 foi que se pudesse categorizar corretamente os autos dos processos da Vara de Execução Fiscal do Distrito Federal, para que a digitalização processual ocorresse corretamente. A Figura 8 ilustra um exemplo de processo digitalizado com seus anexos identificados corretamente.

Figura 8 - Exemplo de processo digitalizado com documentos anexados corretamente



Nesse sentido, o projeto do TJDFT apresenta uma característica inovadora, pois foi utilizada uma aprendizagem não supervisionada. Ou seja, um algoritmo de *cluster* (agrupamento) foi submetido a um conjunto de diferentes tipos de arquivos processuais, sem saber exatamente quais tipos de documentos estavam sendo analisados. Portanto, não houve a necessidade de um especialista para mapear os documentos previamente.

Uma premissa importante é a de que uma massa significativa de documentos seja disponibilizada para o treinamento da máquina. Esse treinamento não supervisionado força o Hórus a tentar encontrar padrões e relacionamentos entre os documentos para que seus arquivos sejam divididos em categorias, de acordo com o tipo (Mandado, Certidão, Petição, etc.). Quanto mais dados/documentos disponíveis, mais eficaz é o treinamento.

A partir do treinamento inicial de descoberta de padrões, armazenamos informações importantes, como:

- 3.3.1.1 Quantidade de grupos formados;
- 3.3.1.2 Tipo de documento de cada grupo;
- 3.3.1.3 Parâmetros de configuração/calibragem do treinamento.

A informação do item 3 listada anteriormente é muito importante para os experimentos, para que se possa sempre comparar os valores dos parâmetros escolhidos até que se atinja um bom nível de precisão nos resultados.

3.3.2 TRATAMENTO DO DOCUMENTO TEXTUAL

Após a etapa descrita na seção "Digitalização e OCR", temos como material inicial para a tarefa de classificação do Hórus os autos processuais em formato de arquivos de texto. O primeiro passo é tratarmos esses arquivos para que possam ser utilizados no treinamento do Hórus.

O ideal é não trabalharmos com os arquivos textuais puros, mas convertê-los em um formato compreensível pela máquina; neste caso, em um conjunto de características principais do documento. É a partir desses conjuntos de

características que os padrões e relacionamentos serão identificados e os agrupamentos formados para cada tipo de documento.

As características de um auto processual utilizadas no Hórus baseiam-se em seu conteúdo, contrastando também com o conteúdo dos outros autos utilizados para treinamento. Em outras palavras, utilizamos uma medida estatística cujo objetivo é indicar a importância das palavras em um documento com relação a uma coleção de documentos. Essa medida é conhecida como TF-IDF (*Term Frequency-Inverse Document Frequency*), muito utilizada em mineração de textos e recuperação de informação (RAJAMARAN; ULLMAN, 2011).

Basicamente, o valor TF-IDF aumenta proporcionalmente ao número de vezes em que uma palavra aparece no documento, porém, é equilibrado pela frequência da palavra no conjunto total de documentos. A ideia central é a de que se uma palavra aparece muitas vezes em um documento, sua relevância deve ser considerada, pois deve ser uma palavra mais significativa do que outras que aparecem poucas vezes.

Ao mesmo tempo, se essa palavra aparece poucas vezes em outros documentos, deve ser uma palavra representativa de um tipo específico de auto processual. Caso essa palavra apareça muitas vezes em todos os outros documentos, não deve ser um diferencial e provavelmente será ignorada no nosso conjunto de características, pois não será capaz de representar um tipo de documento específico.

Em resumo, cada arquivo de texto será transformado em um conjunto de características baseadas nas palavras presentes em seus conteúdos e nos conteúdos dos outros documentos. Após o cálculo do TF-IDF, o valor é comparado a um limiar, configurado pelo desenvolvedor, para saber se a palavra será considerada para compor o conjunto de características do documento.

Outras tarefas também fazem parte do tratamento do conteúdo dos autos para a formação das suas características:

- *Stemming* (LOVINS, 1968);
- Eliminação de *stopwords* (RAJAMARAN; ULLMAN, 2011).

O *Stemming* é o processo de reduzir palavras ao seu radical (ou base, tronco, *stem*). Deve ser uma redução suficiente para que palavras relacionadas sejam mapeadas para a

mesma base. Por exemplo, se possuímos as palavras "certificado" e "certificar" em um documento, levaremos em conta a mesma base ("certif", por exemplo) para a formação das características desse documento, eliminando assim redundâncias no processamento.

A outra tarefa é a eliminação de *stopwords*. *Stopwords* são palavras que, de maneira geral, não representam informações relevantes para o documento, como artigos, preposições, ou pronomes. Esses termos não devem então ser considerados na formação das características dos

autos.

No Hórus, essas duas tarefas são configuráveis:

- Podemos escolher se o tratamento dos documentos será executado com ou sem *Stemming*;
- A lista de *stopwords* é editável, ou seja, sempre podemos adicionar novas palavras que não devem ser consideradas no tratamento dos arquivos textuais dos autos.

3.3.3 TREINAMENTO E CLASSIFICAÇÃO

Nesse ponto, mais uma etapa foi executada pelo Hórus: a preparação dos documentos textuais. A saída dessa etapa, conforme descrito anteriormente (TF-IDF, *Stemming*, eliminação de *Stopwords*), é apresentada como conjuntos de características, contendo palavras principais (neste caso, com *Stemming*) e pesos associados (valores TF-IDF), como podemos observar na Tabela 1.

Tabela 1 - Exemplo de representação de características de um auto processual

```
[andament=3.7969462747702254,assunt=3.7969462747702254,ativ=0.5151285085250104,
citaca=0.5151285085250104,class=0.654705218250406,determin=0.5198603854199589,
execuca=0.674337848602449,exequent=0.6836663453462406,expedica=4.316806660190184,
federa=4.316806660190184,fiscal=0.674337848602449,fpdf=0.6490448140238687,inclu=3.277085889350266,
juiz=0.501109408765896,port=0.7015954859864064,process=1.3533738712541137,territori=0.5198603854199589,
tjdf=3.7969462747702254,transport=3.109728225864609,tribunal=0.5198603854199589,unia=3.277085889350266,
var=0.5057533864518368]
```

Todas essas representações dos documentos são utilizadas para a aprendizagem não supervisionada do Hórus. São essas representações que são agrupadas de acordo com suas semelhanças.

A aprendizagem utiliza o clássico algoritmo *K-Means* (MACQUEEN *et al*, 1967) para o processo de *clustering*. Em mineração de dados, agrupamento *K-Means* é um método que possui como objetivo dividir *n* observações (no nosso caso, documentos) dentre *k* grupos.

Cada documento deve pertencer ao grupo mais próximo à média das suas características.

No Hórus, a Inteligência Artificial é explorada em duas fases:

- 3.3.3.1 Treinamento;
- 3.3.3.2 Classificação.

O treinamento consiste em, dado um conjunto de autos processuais, aprender a dividi-los em grupos, de acordo com o tipo de auto (Certidão, Petição, Mandado, etc.). O *K-means* é um algoritmo iterativo, baseado em centróides (núcleos). A escolha da quantidade de grupos (*k*) é feita previamente. Na primeira iteração, os documentos iniciais são postos aleatoriamente nos grupos, enquanto os documentos restantes são colocados nos grupos com os quais possuem maior similaridade. Ao calcularmos uma

média das características dos documentos de cada *cluster*, temos o núcleo do grupo, também representado em função de palavras e pesos.

Nas iterações seguintes, tentamos reposicionar os documentos com base nos centróides calculados de cada *cluster*, sempre levando em conta a similaridade. Com isso, teremos grupos com configurações diferentes, consequentemente novos núcleos, mais próximos do ideal do que na iteração anterior.

As iterações terminam quando (i) o número máximo de repetições é alcançado; ou

(ii) nenhum documento move-se para outro grupo. Esse número máximo de iterações também é configurado pelo desenvolvedor. Todos os parâmetros configuráveis são importantes para que se explore a melhor combinação de valores nos experimentos, que levem ao melhor resultado possível.

Quando o método encerra, temos os grupos formados e seus documentos, assim como os núcleos calculados de cada grupo (representação idêntica à ilustrada na Tabela 1). Essas e outras informações são persistidas em uma base de dados após cada treinamento do Hórus, conforme estrutura mostrada na Figura 9.

Figura 9 - Estrutura da tabela de treinamento do Hórus

COLUMN_NAME	DATA_TYPE	NULL...	DATA_DEFAULT	COLUM...	COMMENTS
1 ID	NUMBER	No	"HORUS"."ISEQ\$\$_2845331".nextval	1	(null)
2 NUMCLUSTER	NUMBER(10,0)	No	(null)	2	(null)
3 CENTROIDE	VARCHAR2(5000 BYTE)	No	(null)	3	(null)
4 LOCALIZACAO	VARCHAR2(100 BYTE)	No	(null)	4	(null)
5 TIPODOCUMENTO	VARCHAR2(100 BYTE)	No	(null)	5	(null)
6 MINSIGTFIDF	NUMBER(10,2)	No	(null)	6	(null)
7 CUTOFF	NUMBER(10,0)	No	(null)	7	(null)
8 ITERACOESKMEANS	NUMBER(10,0)	No	(null)	8	(null)
9 USOUSTEMMER	NUMBER(1,0)	No	(null)	9	(null)
10 DATA_HORA	TIMESTAMP(6)	Yes	CURRENT_TIMESTAMP	10	(null)

A classificação consiste em, dado um auto processual específico de um dos tipos utilizados na fase de treinamento, informar a que grupo ele pertence, ou seja, informar o seu tipo, baseado também em seu conteúdo. Para isso, recuperamos os núcleos persistidos no Banco de Dados do nosso treinamento mais recente. É a partir desses núcleos que é calculada a similaridade do novo auto, representado por suas características, assim como na fase de treinamento. No caso da fase de classificação, apenas uma iteração do K-Means é executada.

3.4 ENVIO AO PJE

O PJE é o sistema de tramitação processual mantido pelo CNJ para processos de 1ª e 2ª instâncias, entretanto customizações e novas implementações podem ser feitas por cada tribunal individualmente. Atualmente, o TJDFT é o maior parceiro de desenvolvimento dessa solução nacional de controle de processos judiciais. Dentre as diferentes tecnologias que compõem essa aplicação, podemos destacar a interface MNI, que possibilita a comunicação com outras instituições. Essa interface permitiu que os processos digitalizados pudessem ser importados para o PJE sem a necessidade de cadastros manuais.

Algumas particularidades do PJE requerem alguns tratamentos especiais, como a assinatura de todos os documentos que são carregados, utilizando um padrão de certificação digital reconhecido pela ICP-Brasil; a identificação de todos os tipos de documentos baseados em códigos específicos do PJE; a presença de informações de identificação dos pólos processuais com os respectivos detalhes do endereço de pessoa física ou jurídica.

4 RESULTADOS

O projeto Hórus encontra-se em processo de homologação por parte da VEF. A previsão é de que seja posto em produção nos primeiros dias do mês de maio de 2019, com a meta inicial de serem distribuídos 48.000 processos já digitalizados, indexados e fragmentados de forma automatizada e inteligente. Durante a fase de testes, foram realizadas simulações com diferentes tipos de processos e várias situações contribuíram para a melhoria do Hórus, com destaque para:

Testes com 2 núcleos de classificação (Certidão e CDA), tendo como acurácia 92% de acerto com um volume de treinamento de pouco mais de 800 arquivos textuais;

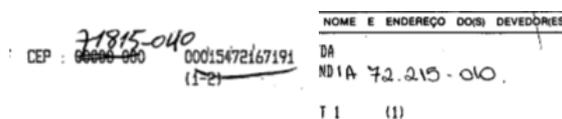
Contabilização de 8 segundos para a execução de todas as fases do Hórus para um processo com características

comuns, isto é, uma certidão e uma CDA, que configura cerca de 90% do universo de processos da VEF;

Distribuição automática de processos, evitando a necessidade destes serem distribuídos por um centro especializado nessa atividade, a qual foi identificada como uma atividade morosa, repetitiva e desgastante.

Informações de endereço da parte que não estão presentes no cadastro dos sistemas legados. Estas são complementadas com as existentes na CDA, a fim de qualificar o dado a ser enviado ao PJE, evitando uma atualização do processo após a distribuição. Contudo, algumas informações manuscritas dificultam essa atividade complementar, como pode ser observado na Figura 10.

Figura 10 - Dados manuscritos não processados pelo Hórus



5 CONCLUSÃO

Em detrimento das particularidades existentes na VEF, o projeto tem apresentado resultados satisfatórios e promissores com relação ao uso de Inteligência Artificial em suas atividades diárias, devido ao volume de trabalho e padronização de algumas informações básicas de atuação do processo, as quais são oriundas de um formulário padronizado encaminhado pela Fazenda Pública do Distrito Federal (FPDF).

A morosidade na distribuição dos processos da unidade pelo órgão especializado do TJDFT foi solucionada pelo Hórus com a distribuição automática, onde o órgão julgador e o número de identificação do processo junto ao CNJ existentes no legado são diretamente sensibilizados na carga de dados ao PJE, preservando sua localização e identificação, além de solucionar um problema que vem se agravando com a determinação da administração do TJDFT em finalizar o biênio com boa parte do seu acervo de processos em plataforma digital. Especificamente com relação à classificação dos documentos do processo, que efetivamente incorporou tecnologias de aprendizado de máquina, buscou-se complementar as atividades de fragmentação e indexação, analisando os documentos do processo e verificando se o tipo fragmentado realmente atendia às especificações do treinamento, isto é, foi realizada uma conferência do tipo fragmentado com o tipo

classificado. Caso houvesse divergência na classificação, o tipo fragmentado era considerado como padrão, porém o respectivo arquivo foi identificado e destacado para posterior melhoria das regras de classificação.

É notória a relevância da parceria entre a área do Direito e a da Tecnologia da Informação, a fim de que o jurisdicionado e o governo do Distrito Federal possam usufruir de um sistema robusto, eficaz e eficiente.

REFERÊNCIAS

BRASIL. Lei nº 6.830, de 22 de setembro de 1980. Dispõe sobre a cobrança judicial da Dívida Ativa da Fazenda Pública, e dá outras providências. In: Diário Oficial da União, Brasília, DF, 24 set. 1980. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l6830.htm>. Acesso em: 20 mar. 2019

BRASIL. Lei nº 11.697, de 13 de junho de 2008. Dispõe sobre a organização judiciária do Distrito Federal e dos Territórios e revoga as Leis nos 6.750, de 10 de dezembro de 1979, 8.185, de 14 de maio de 1991, 8.407, de 10 de janeiro de 1992, e 10.801, de 10 de dezembro de 2003, exceto na parte em que instituíram e regularam o funcionamento dos serviços notariais e de registro no Distrito Federal. In: Diário Oficial da União, Brasília, DF, 16 jun. 2008. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2008/Lei/L11697.htm>. Acesso em: 20 mar. 2019

Conselho Nacional de Justiça - CNJ. Justiça em Números 2018. Disponível em: <<https://www.cnj.jus.br/files/conteudo/arquivo/2018/09/8d9faee7812d35a58cee3d92d2df2f25.pdf>>. Acesso em: 20 mar. 2019.

HECKSTER, Nicky. 5 Predictions in 5 years: What do *Big data* Environment Will Look Like. *Big data* Expo 2015. IBM. Disponível em: <<https://www.slideshare.net/BigDataExpo/big-data-expo-2015-ibm-5-predictions>>. Acesso em: 29 mar. 2019.

LOVINS, Julie Beth. Development of a *Stemming* algorithm. *Mech. Translat. & Comp. Linguistics*, v. 11, n. 1-2, p. 22-31, 1968.

MACQUEEN, James *et al.* Some Methods for classification and analysis of multivariate observations. In: *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*. 1967. p. 281-297.

RAJAMARAN, Anand; ULLMAN, Jeffrey David. Mining of massive datasets. Cambridge University Press, 2011. RICH, Elaine *et al.* Inteligência artificial. 1994.

SHANNON, Robert E.; MAYER, Richard; ADELSBERGER, Heimo H. Expert systems and simulation. *Simulation*, v. 44, n. 6, p. 275-284, 1985.

SMITH, Ray. An overview of the *Tesseract* OCR engine. In: *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*. IEEE, 2007. p. 629- 633.

Supremo Tribunal Federal. Notícias STF, de 30 de maio de 2018. Inteligência Artificial vai agilizar a tramitação

de processos no STF. Disponível em: <<http://www.stf.jus.br/portal/cms/verNoticiaDetalhe.asp?idConteudo=380038>>. Acesso em: 20 mar. 2019.

Superior Tribunal de Justiça. Notícias, de 19 de fevereiro de 2009. Projeto de digitalização dos processos no STJ mobiliza tribunais. Disponível em: <<https://stj.jusbrasil.com.br/noticias/810630/projeto-de-digitalizacao-dos-processos-no-stj-mobiliza-tribunais>>. Acesso em: 29 mar. 2019.

Tribunal de Justiça de Pernambuco. Notícias, de 11 de março de 2019. Integrantes do TJPE apresentam ferramenta de Inteligência Artificial em eventos. Disponível em: <http://www.tjpe.jus.br/noticias/-/asset_publisher/ubhL04hQXv5n/content/integrantes-do-tjpe-apresentam-ferramenta-de-inteligencia-artificial-em-eventos?inheritRedirect=false&redirect=http%3A%2F%2Fwww.tjpe.jus.br%2Fnoticias%3Fp_p_id%3D101_INSTANCE_ubhL04hQXv5n%26p_p_lifecycle%3D0%26p_p_state%3Dnormal%26p_p_mode%3Dview%26p_p_col_id%3Dcolumn-2%26p_p_col_count%3D1>. Acesso em: 20 mar. 2019.

Tribunal de Justiça de Rondônia. Notícias, de 07 de Junho de 2018. Inteligência Artificial desenvolvida pelo TJRO pode revolucionar o Judiciário. Disponível em: <<https://www.tjro.jus.br/noticias/item/9472-inteligencia-artificial-desenvolvida-pelo-tjro-pode-revolucionar-o-judiciario>>. Acesso em: 20 mar. 2019.

TRIBUNAL DE JUSTIÇA DO DISTRITO FEDERAL E DOS TERRITÓRIOS.

Resolução do Pleno Administrativo n. 19, de 2 de dezembro de 2009. Dispõe sobre a instalação da Vara de Execução Fiscal no Distrito Federal. Disponível em: <<https://www.tjdft.jus.br/publicacoes/publicacoes-oficiais/resolucoes-do-pleno/2009/00019.html>>. Acesso em: 20 mar. 2019

TRIBUNAL DE JUSTIÇA DO DISTRITO FEDERAL E DOS TERRITÓRIOS.

Provimento da Corregedoria n. 13, de 9 de outubro de 2012. Dispõe sobre o arquivamento sem baixa das execuções fiscais com valores inferiores a R\$ 5.000,00 (cinco mil reais). Disponível em: <<https://www.tjdft.jus.br/publicacoes/publicacoes-oficiais/provimentos/2012/13>>. Acesso em: 20 mar. 2019

TURING, Alan M. Computing machinery and intelligence. In: *Mind*, v. 59, n. 236, p. 433, 1950.

VAZ, Conrado Adolpho. *Google Marketing-O Guia Definitivo de Marketing Digital*. Novatec Editora, 2009.

VISWARUPAN, Niruhan. *K-Means Data Clustering*. Disponível em: <<https://towardsdatascience.com/K-Means-data-clustering-bce3335d2203>>. Acesso em: 29 mar. 2019.

WATERMAN, Donald. *A guide to Expert Systems*. 1986.

Dr. Jairo Simão Santana Melo
Supervisor e Colaborador do Serviço de Gestão de Dados do TJDFT.

Dr. Weiss Webber Araújo Cavalcante
Juiz Titular da Vara de Execução Fiscal do TJDFT.