

Inteligência artificial aplicada para o aumento da produtividade no atendimento de intimações

Marcus Parreiras², Antônio Vasconcellos³, Eduardo Mangeli¹, Emerson Yamamoto², Geraldo Xexéo², Izandro Metello², Lincoln Costa², Pedro Marques¹, Jano de Souza²

¹ Instituto de Computação
Universidade Federal do Rio de Janeiro - Rio de Janeiro - Brasil

² Programa de Eng. de Sist. e Computação, COPPE
Universidade Federal do Rio de Janeiro - Rio de Janeiro - Brasil

³ PUC-Rio - Rio de Janeiro - Brasil

Abstract. *This paper describes an initiative to use Artificial Intelligence within the scope of the Rio de Janeiro State Public Defender's Office, the CICIAR. This work presents the motivation for its development and the benefits of the users of the implemented service. The results point to improvements for the Rio de Janeiro State Public Defender's Office, both in streamlining the workflow and in workload planning, and consequently for the population that uses it, such as an increase in speed and greater accuracy of the service.*

Resumo. *Este artigo apresenta uma iniciativa de utilização de Inteligência Artificial no âmbito da Defensoria Pública do Estado do Rio de Janeiro, a Caixa de Intimação com Inteligência Artificial (CICIAR), desde a motivação para o seu desenvolvimento até os benefícios dos usuários do serviço implementado. Os resultados apontam melhorias para a Defensoria Pública do Estado do Rio de Janeiro, tanto na agilização do fluxo de trabalho como na melhor distribuição e planejamento do trabalho, e consequentemente para a população que a utiliza, como aumento na celeridade e maior acuidade do serviço.*

1. Introdução

Segundo Pedruzzi (2020), a combinação de máquinas e processos digitais capazes de tomar decisões são vetores de uma mudança de paradigma na forma de pensar o Direito. Como consequência inicial, a prática da busca de informações evoluiu de consultas a livros, revistas de Direito, e legislação em cópia impressa, para o uso de bibliotecas digitais. Em seguida houve uma informatização de escritórios de advocacia e departamentos jurídicos. Órgãos governamentais do Brasil também estão migrando acervos para o meio digital como parte dessa transição. Uma terceira etapa surge com a utilização de técnicas de Inteligência Artificial (IA), que possibilitam a busca de informações de forma mais rápida com menos custo e maior precisão [Pedruzzi 2020]

Neste contexto, o CICIAR – Caixa de Intimações com Inteligência Artificial – foi criado para auxiliar defensores públicos a tratar e priorizar o atendimento das intimações recebidas no Verde, plataforma digital usada pela Defensoria Pública do Estado do Rio de Janeiro (DPRJ) para gerenciar o acompanhamento de processos. O CICIAR tem como diferencial categorizar as intimações recebidas sem acessar seu conteúdo, o que iniciaria a

contagem do prazo para o seu atendimento. Essa categorização auxilia o planejamento do defensor, sem registrar a leitura do documento de intimação, ou visualizar seu conteúdo.

O sistema representa a consolidação do conhecimento de especialistas no domínio do problema, defensores e servidores que participaram das etapas iniciais de elicitação de conceitos de interesse, e heurísticas capazes de identificar categorias de intimações, além da construção de uma base de dados históricos de intimações com a respectiva categorização indicada. Ainda, durante o piloto e após a implantação do CICIAR, ao operar do sistema, os usuários podem indicar a qualidade dos resultados, ao confirmar ou corrigir as categorias atribuídas automaticamente às intimações.

O projeto CICIAR deve ser entendido no contexto do planejamento da DPRJ de aplicação de tecnologia, em especial IA, para impactar positivamente na produtividade do órgão e, conseqüentemente, melhorar o atendimento à comunidade.

Entre as motivações para o uso de IA na DPRJ, podemos destacar:

- Busca por melhoria na gestão do órgão;
- Possibilidade de estruturação de fluxos e processos de trabalho;
- Melhor distribuição da carga de trabalho;
- Possibilidade de antecipação das intimações e preparação das atividades decorrentes; e
- Possibilidade de planejamento estratégico de longo prazo.

A plataforma digital Verde, utilizada pela DPRJ, registrou o recebimento de, aproximadamente, 107.000 avisos de intimação por mês durante o ano de 2021. Este grande volume de intimações, aliado à complexidade natural inerente às ações necessárias para seu tratamento, motivou a DPRJ a investigar o desenvolvimento de ferramentas de IA capazes de auxiliar os defensores e defensoras nesse trabalho. O projeto CICIAR surge dessa iniciativa.

Atualmente o CICIAR é uma funcionalidade ativa e em uso por todos os órgãos que usam a funcionalidade *Caixa de Intimações* do Verde. O objetivo deste artigo é descrever o CICIAR e quais os impactos iniciais do seu uso, do ponto de vista do relato dos defensores. Para isso, o artigo está organizado da seguinte forma:

- Na seção 2 descrevemos o referencial teórico.
- Na seção 3 apresentamos outras iniciativas e trabalhos relacionados ao tema.
- Na seção 4, introduzimos o processo de implementação do sistema e seus detalhes estratégicos.
- Na seção 5, descrevemos a arquitetura do CICIAR
- Por fim, na seção 6 concluímos o trabalho descrevendo o impacto do sistema e nossos próximos passos.

2. Referencial Teórico

Esta seção apresenta o estado da arte dos temas abordados, sob aspectos teóricos, e estudos anteriores de aplicações de IA no governo e soluções de *Deep Learning* e Processamento de Linguagem Natural.

2.1. Inteligência Artificial e o Governo

Silva et. al afirmam que “mesmo em cenários de forte automação, os seres humanos ainda serão necessários na força de trabalho e os ganhos totais de produtividade que são estimados pelos estudos só acontecerão se as pessoas trabalharem lado a lado com as máquinas” [da Silva et al. 2021]. Os autores também consideram que o uso da Inteligência Artificial aplicada ao governo vai mudar profundamente o seu funcionamento, afetando áreas como planejamento e gestão de recursos[da Silva et al. 2021].

Alguns autores manifestam preocupação com possíveis problemas que o uso de Inteligências Artificiais aplicadas a governos possam causar. Divino e Magalhães [Divino and Magalhães 2020] citam que “a observância dos princípios éticos da autonomia, prevenção de danos, equidade, explicabilidade deve ser realizada para tornar possível a compatibilização do desenvolvimento tecnológico com as premissas econômicas e a tutela jurídica de direitos individuais”.

Mehr [Mehr et al. 2017] apresenta uma pesquisa sobre a atuação de inteligências artificiais aplicadas à governos, e aponta que o setor público vem demonstrando um crescente interesse no uso desse tipo de abordagem. O trabalho ressalta que existem pouquíssimos estudos neste sentido. Apesar disso, afirma que Inteligência Artificial aplicada no setor público tem vantagens óbvias, e lista “redução de trabalhos administrativos, ajuda na alocação de recursos e substituição de humanos em tarefas muito complexas” [Mehr et al. 2017].

Outras vantagens para o setor público incluem, mas não se limitam a [Chatfield and Reddick 2019]:

- Aumento da eficiência de trabalho;
- Diminuição no tempo para o processamento das requisições;
- Redução da carga de trabalho humano;
- Promoção do crescimento econômico via redução de gastos; e
- Melhoria generalizada no fluxo de processo do trabalho e redução de erros.

Silva et. al [da Silva Duarte et al. 2020] apresentam uma pesquisa sobre o uso de inteligências artificiais por governos pelo mundo, sobretudo em programas com altos investimentos nesta abordagem.

Nesse sentido, o governo da China pretende investir, até 2030, um total de 147,8 bilhões de dólares americanos, com o objetivo de se tornar um líder mundial em inovações em Inteligência Artificial, e os Estados Unidos investiram cerca de 1,2 bilhões de dólares americanos apenas em 2016 em pesquisa e desenvolvimento de inteligências artificiais, tendo uma abordagem relacionada ao inventivo de pesquisa, educação e especialização na área a longo prazo [Bundy 2017].

2.2. Deep learning

De acordo com LeCun et al. e Bengio ([LeCun et al. 2015] e [Bengio 2009]), o aprendizado profundo e neural são técnicas bem estabelecidas no campo de aprendizado de máquina e têm sido bem sucedidos em uma variedade de tarefas como processamento de imagem e texto [Goldberg 2016], [Bengio et al. 2000], [Mikolov et al. 2013], [Collobert and Weston 2008], [Collobert et al. 2011], [Socher et al. 2013], [Zhang and Wallace 2015].

Segundo LeCun et al. [LeCun et al. 2015], técnicas convencionais de aprendizado de máquina eram limitadas na capacidade de processar dados naturais em sua forma bruta, exigindo uma engenharia cuidadosa e grande conhecimento do domínio para construir extratores de atributos que transformam dados brutos, como pontos de uma imagem, em uma representação adequada para o sistema de aprendizado. Métodos de aprendizado profundo permitem a entrada de dados brutos e automaticamente descobrem representações necessárias para detecção ou classificação.

2.3. Processamento de Linguagem Natural de texto

Rao e McMahan [Rao and McMahan 2019] afirmam que o Processamento de Linguagem Natural se refere a um conjunto de técnicas que envolvem a aplicação de métodos estatísticos, com ou sem o apoio da Linguística, para entender texto com o objetivo de resolver problemas do mundo real. Essa compreensão vem principalmente da transformação de textos em representações computacionalmente úteis, como estruturas de vetores, grafos ou árvores.

Dados textuais podem ter origem de diversas fontes, como páginas web, e-mails, redes sociais e documentos jurídicos, para citar algumas. A extração de informações de textos pode ser desafiadora, dada a sua natureza desestruturada.

3. Trabalhos Relacionados

Esta seção destaca os trabalhos relacionados, a fim de mostrar o que já foi feito de importante e resultados obtidos sobre aplicações parecidas, de modo que seja possível complementar ou comparar o mostrado neste trabalho.

O PIÁ (Paraná Inteligência Artificial) é um portal do governo do estado do Paraná que foi lançado em 2019 pela Companhia de Informação e Comunicação do Paraná (CELEPAR), em Foz do Iguaçu, com a proposta de tornar o estado conhecido como o mais inovador do Brasil [da Silva Duarte et al. 2020].

O objetivo do PIÁ é tornar mais ágeis, baratas e transparentes movimentações do serviço público. Através de um cadastro único, utilizando documentação como carteira de motorista ou RG, o cidadão pode “solicitar a renovação de documentos pessoais, agendar atendimento em clínicas e centros de saúde e obter faturas de energia elétrica e de água, seja utilizando o celular, o tablet ou o computador” [da Silva Duarte et al. 2020]. O nome “Piá” é um vocábulo paranaense sinônimo de “menino”.

Uma solução de instrução assistida para elaboração de pareceres sobre processos judiciais [Dutra et al. 2021] foi projetada para auxiliar o trabalho diário de servidores do Tribunal de Contas da União. A Inteligência Artificial nessa aplicação representa o conhecimento do servidor, que busca manualmente as informações em repositórios de teses ou arquivos da Conjur. Como resultado, os ganhos de produtividade para processos repetitivos foram de 70% a 80%, um trabalho que demoraria de 6 a 8 horas foi reduzido para 1 hora. Para processos de maior complexidade, o incremento de produtividade ficou entre 20% e 50%.

A área de Processamento de Linguagem Natural em Direito carece de recursos textuais [Martim et al. 2018]. Visando auxiliar na solução desse problema, [Sousa and Del Fabro 2019] construíram uma base com mais de 50 mil documentos

jurídicos produzidos em um intervalo de 9 anos advindos de acórdãos do Supremo Tribunal Federal. Essa base busca fornecer recursos para as mais variadas tarefas da área de NLP, em contraste com bases menores voltadas para tarefas específicas, como Reconhecimento de Entidades Nomeadas (NER).

Apesar da escassez de recursos textuais no idioma Português, pode-se encontrar aplicações de técnicas de NLP como modelagem de tópicos [de Souza Tavares et al. 2020], classificação [Dal Pont et al. 2021], resumo automático [de Sousa and Prata 2019] e predição [Bertalan 2020].

4. Implementação

O passo inicial da implementação envolveu a investigação e documentação de como o processo de identificação dos temas das intimações era realizado pelos defensores. Em cooperação com representantes da DPRJ, a equipe CICIAR mapeou as principais etapas e elementos envolvidos, de forma que fosse construída uma estratégia de solução integrada à plataforma Verde. Parte desta estratégia vai além da abordagem do problema, como a organização da equipe e as ferramentas utilizadas.

4.1. Verde: o Sistema Facilitador da DPRJ

Visando aprimorar e facilitar o acesso dos cidadãos assistidos pela DPRJ ao agendamento dos atendimentos e ao trabalho de cada órgão da Defensoria, no dia 1º de junho de 2016 foi lançado o Verde.

Com 4.822 usuários ativos e mais de 300 mil casos abertos somente em 2021, o Verde contempla os principais fluxos de trabalho executados por um órgão da Defensoria durante a sua rotina: cadastro unificado de assistidos, cadastro de casos, emissão de documentos e ofícios, cadastro de lembretes, visualização completa da agenda, recebimento e gerenciamento de intimações, entre outros.

O grande fluxo de dados gerado pelos assistidos e seus processos faz com que apareça um cenário propício para a aplicação de técnicas de Inteligência Artificial, e é nesse contexto que surge o CICIAR, inicialmente visando categorizar as intimações recebidas pelos defensores públicos antes mesmo que estas sejam abertas.

4.2. Organização da Equipe

A equipe CICIAR utiliza no desenvolvimento de suas atividades a metodologia ágil *Scrum* [Schwaber and Sutherland 2011]. Uma das suas principais características é a separação das etapas do desenvolvimento do projeto em intervalos de tempo fixos denominados *sprints*, finalizadas sempre com a reuniões de entrega e apresentação de funcionalidades novas para as partes interessadas. Essas reuniões são marcos essenciais de validação para a equipe, que de forma iterativa melhora seu entendimento do problema em solução. No caso particular do projeto, foram utilizadas *sprints* quinzenais.

Além disso, são realizadas reuniões diárias e rápidas (chamadas *dailies*) entre os membros da equipe CICIAR, que permitem o relato do progresso realizado por cada um e dos obstáculos encontrados. Ao longo da reunião, um *Kanban* é utilizado para a visualização do estado das tarefas e o andamento das metas da *sprint*.

4.3. Ferramentas Utilizadas

A plataforma Gitlab¹ é utilizada para integrar o versionamento do código com as funcionalidades de gestão ágil, como o *Kanban* mencionado.

Para experimentos de análise de dados a equipe utiliza *Jupyter Notebooks*², que facilitam a geração de relatórios. Para experimentos envolvendo treinamento de modelos de IA e outras atividades computacionalmente custosas, o *Google Colab*³ fornece uma infraestrutura com GPU integrada, tornando mais eficiente e veloz as execuções.

4.4. Construção de uma Base de Dados

Em paralelo ao mapeamento de etapas e elementos do processo manual dos defensores, a equipe construiu sua base de dados utilizando um Data Warehouse. A modelagem do banco de dados foi feita a partir do estudo das estruturas do Verde e do MNI⁴. As fontes utilizadas para extração foram os bancos de dados da plataforma Verde e do Conselho Nacional de Justiça (CNJ), além do Web Service do Tribunal de Justiça (TJ). Os dados foram importados para a base a partir de um processo de ETL (do inglês *Extract, Transform, Load*) que consiste na extração, tratamento e carregamento de dados.

Com a base pronta, foram realizados experimentos de verificação de integridade e análises de dados.

4.5. Estratégia de Solução

De acordo com os relatos dos representantes da DPRJ, na maioria dos casos o processo manual de descobrir os temas das intimações consiste na leitura direcionada dos andamentos mais recentes do processo, sem ler a intimação. Os movimentos lidos seriam os recebidos entre o aviso e a última decisão do juiz. Este subconjunto foi denominado Contexto de um Aviso (ou apenas Contexto). Este objeto e a abordagem para sua análise compõe o núcleo do sistema CICIAR.

Para automatizar a análise de Contextos e descobrir seus temas, a equipe estabeleceu o problema como um caso de classificação e utilizou ferramentas das áreas de Aprendizado por Máquina e Redes Neurais.

Como uma intimação pode conter mais de um tema, classificar seu Contexto é, a princípio, um problema multi-rótulo, em que cada aviso poderia ser classificado como mais de uma classe. É dito “a princípio”, pois é possível transformar o problema em um caso mais simples de classificação. A abordagem consiste em dividir um contexto em múltiplas frases, tal que cada uma pode ter apenas um rótulo atribuído. Dessa forma, para definir os temas de um Contexto, é necessário classificar todas as suas frases, obtendo um conjunto de temas detectados. Essa divisão em frases é feita por extração de documentos e etapas de processamento de texto. Os rótulos estão listados na tabela Tabela 1.

Para transformar as frases em objetos interpretáveis matematicamente e manipuláveis computacionalmente, foi utilizado o BERTimbau [Souza et al. 2020], um modelo pré-treinado de processamento de texto, que pode ser utilizado para atribuir para

¹<https://gitlab.com/>

²<https://jupyter.org/>

³<https://colab.research.google.com/>

⁴<http://www.tjrj.jus.br/web/guest/intrav2/manuais/manuais/manuais-e-videos-publicos/mni-modelo-nacional-de-interoperabilidade>

cada frase um vetor multidimensional numérico, denominado *embedding*. Estes vetores gerados são capazes de representar semelhanças semânticas entre frases.

Para criação de uma base rotulada de frases, importante para treinamento dos modelos de classificação, a equipe desenvolveu uma aplicação web chamada Sistema de Avaliação, que permite que usuários elencados pela DPRJ rotulem as frases de intimações e de seus Contextos.

Tendo os *embeddings* gerados e a base rotulada, modelos foram treinados para atribuir rótulos temáticos (Tabela 1) a cada frase. O CICIAR utiliza dois classificadores de redes neurais em sua arquitetura:

- **Classificador Binário:** para diferenciar frases relevantes, que podem sugerir os temas do aviso, de frases irrelevantes, que não indicam tema, por exemplo cabeçalhos de documento.
- **Classificador Multiclasse:** separa as frases relevantes entre os temas possíveis.

Rótulos	Descrição
Em branco	A frase irrelevante.
Nada	A frase é irrelevante, mas pode causar confusão para um leigo.
Outros	A frase é relevante, mas não é compatível com outros rótulos.
Interlocutória	Alimentos Provisórios, Resultado de Penhora, Deferimento/Indeferimento de Tutela de Urgência.
Decisão Final	Acórdão, Sentença, Extinção do Processo, Homologação, Ciência de Sentença.
Prazos / Manifestações	Alegações Finais, Memoriais, Razões de apelação, Contra-razões, Plano de Adjudicação, Réplica.
Prisão	Relaxamento Prisão, Prisão Revogada, Prisão Decretada, Decretação de Prisão, Alvará de Soltura.

Tabela 1. Rótulos atribuídos pelo classificador.

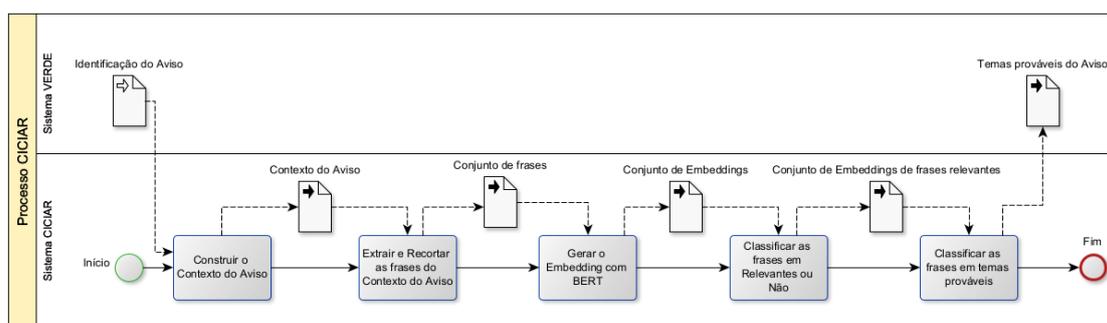


Figura 1. Visão do processo CICIAR

O desenvolvimento do CICIAR depende de ferramentas tecnológicas e de especialistas para fornecer o conhecimento necessário para operar. A quantidade de especialistas e a velocidade de rotulagem são fatores limitantes para o desenvolvimento do sistema.

A partir das etapas apresentadas, o processo do CICIAR foi definido e implementado de forma que recebe de entrada a identificação do Aviso e retorna os possíveis temas do mesmo (Figura 1). O algoritmo funciona continuamente, atendendo uma fila de avisos, que é alterada por meio de um *endpoint* disponibilizado por uma API desenvolvida. Mais detalhes da arquitetura serão detalhados na próxima seção.

5. Arquitetura

No entorno do algoritmo CICIAR estão presentes diversos sistemas que se inter relacionam, gerando um ecossistema complexo. Nesta seção serão apresentados dois sistemas: o primeiro relacionado ao funcionamento diário do CICIAR e o segundo relacionado ao processo de aprendizagem e experimentação.

Além de uma descrição textual, ambos os ecossistemas serão ilustrados utilizando diagramas em Archimate [Josey et al. 2016], uma linguagem de modelagem de arquiteturas. A partir dos símbolos e cores, é possível identificar as tecnologias, componentes, interfaces, processos e funcionalidades.

5.1. API CICIAR : funcionamento diário

A figura 2 ilustra a arquitetura do sistema CICIAR com enfoque na operação diária. O CICIAR executa internamente diversos processos envolvendo criação do Contexto dos avisos, processamento e balanceamento dos dados, heurísticas, entre outros. A tecnologia utilizada para implementação é a linguagem de programação Python.

A integração do CICIAR com o sistema Verde possui dois elementos principais: o recebimento de avisos para serem classificados pelo CICIAR e o envio dos rótulos descobertos para o Verde. Essa interação ocorre por meio de *endpoints* da API CICIAR, implementada utilizando o módulo FastAPI.

Outra integração é relacionada com o *WebService* do Tribunal de Justiça (TJ), que fornece informações dos metadados do processo. Dessa forma, o CICIAR é capaz de recolher os dados necessários para a classificação dos Avisos.

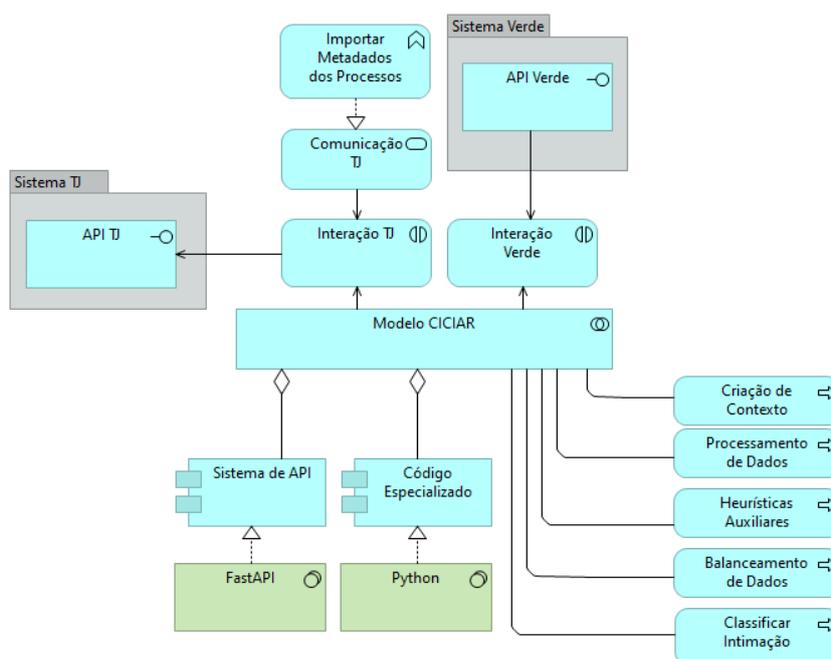


Figura 2. Ecossistema de Funcionamento Diário

5.2. Sistema de Aprendizagem

O Sistema de Aprendizagem do CICIAR possui como principais processos a amostragem dos dados para o treinamento dos modelos, armazenamento das informações dos processos, além da otimização, construção e avaliação dos modelos gerados. Para sustentar esses processos são necessários quatro componentes:

- **ETL:** referente à extração, transformação e armazenamento dos dados. A tecnologia utilizada foi o Pentaho, que insere os dados transformados no Data Warehouse do projeto.
- **Aprendizado:** referente ao treinamento e experimentação de modelos classificadores. A tecnologia utilizada foi o KNIME e o Python.
- **Código Especializado:** este componente agrega funcionalidades internas dos modelos e outros experimentos. Implementações feitas em Python
- **Visualização:** importante componente para análise de dados e tomada de decisões durante a construção de modelos e execução de experimentos. Os gráficos e visualizações são gerados utilizando DataViz.

Os dados que possibilitam o treinamento de modelos e execução de experimentos são adquiridos por meio de três integrações: com o *WebService* do Tribunal de Justiça, que permite acesso aos metadados dos processos; a API do Verde, que fornece dados e o sistema de avaliação, um banco de dados criado pelo próprio CICIAR para armazenamento dos dados para o treinamento do modelo.

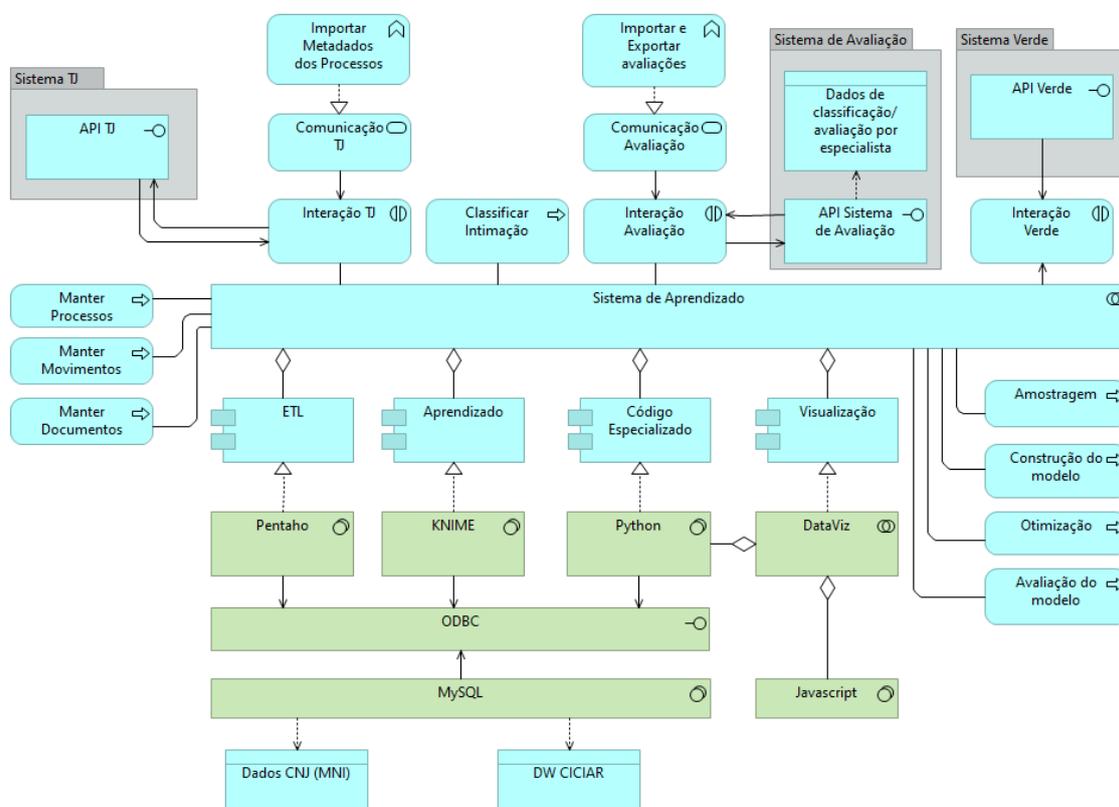


Figura 3. Ecossistema de Aprendizagem

Além disso, o DW também fornece dados para análises e experimentos, que têm

o propósito de extrair informações destes exemplos, que podem ajudar a melhorar o desempenho do CICIAR de modo geral.

6. Conclusão

Esta seção apresenta o fechamento do artigo, e dialoga com os *feedbacks* recebidos até o momento com a utilização da ferramenta.

O CICIAR está implantado em mais de 200 órgãos. A base de dados contém 57397 processos, com um total de 214 mil intimações. Dessas intimações, 68 mil foram destinadas à Defensoria. Até o momento, foram extraídas 27 milhões de frases dos documentos e geradas representações do modelo BERT. Na base, contamos com 42 mil frases classificadas. São classificadas, em média, 2601 intimações por dia. Estão sendo registrados centenas de *feedbacks* semanais dos Defensores em relação às classificações. No ambiente de produção estão habilitados dois rótulos, com previsão de expansão para mais quatro rótulos. Esses números estão em constante mudança, com o aumento da base, melhoria dos modelos, aprimoramento e incremento das heurísticas.

Com base em *feedbacks* de reuniões e nos números de monitoramento da utilização, verificamos que a adesão de utilizadores ao sistema está em constante crescimento. Obtivemos relatos positivos em relação ao auxílio no trabalho dos órgãos quando a carga de trabalho aumenta muito, facilitando a organização das prioridades. Outros pontos destacados são melhorias na gestão, possibilitando criação de fluxos e distribuição do trabalho.

A respostas positivas recebidas após a implantação do CICIAR possibilitaram o desenvolvimento de novas investigações na área de Inteligência Artificial realizadas DPRJ em parceria com a equipe CICIAR. Técnicas de identificação de tópicos estão sendo utilizadas para detecção de conceitos latentes que possam identificar novos rótulos temáticos. Os modelos atualmente em produção estão sendo refinados com o uso de mais exemplos que foram anotados em novas fases de rotulagem feitas por especialistas.

Novas heurísticas candidatas são detectadas, formalizadas e avaliadas constantemente. As heurísticas com bom desempenho são implementadas em produção.

Referências

- Bengio, Y. (2009). *Learning deep architectures for AI*. Now Publishers Inc.
- Bengio, Y., Ducharme, R., and Vincent, P. (2000). A neural probabilistic language model. *Advances in Neural Information Processing Systems*, 13.
- Bertalan, V. G. F. (2020). *Using natural language processing methods to predict judicial outcomes*. PhD thesis, Universidade de São Paulo.
- Bundy, A. (2017). Preparing for the future of artificial intelligence.
- Chatfield, A. T. and Reddick, C. G. (2019). A framework for internet of things-enabled smart government: A case of iot cybersecurity policies and use cases in us federal government. *Government Information Quarterly*, 36(2):346–357.
- Collobert, R. and Weston, J. (2008). A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference on Machine learning*, pages 160–167.

- Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., and Kuksa, P. (2011). Natural language processing (almost) from scratch. *Journal of machine learning research*, 12(ARTICLE):2493–2537.
- da Silva, W. F., Silva, F. S., and da Silva Rabêlo, O. (2021). Tendências no uso de inteligência artificial e sua influência na requalificação da força de trabalho no setor público. *Cadernos de Prospecção*, 14(3):824–824.
- da Silva Duarte, P., Ramos, N. M., Ribeiro, D. Q., and Bastos, A. d. F. V. (2020). A avaliação da acessibilidade web em portais governamentais: o paraná inteligência artificial (piá). *Revista Brasileira de Administração Científica*, 11(1):53–66.
- Dal Pont, T. R., Sabo, I. C., Wilton, P. E. V., Menezes, V. A. d., Copetti, R., Zambrotta, L., Martins, P. P., Costa, E. C., Schnitzler, E. L., Santos, P. M., et al. (2021). Classification and association rules in brazilian supreme court judgments on pre-trial detention. In *International Conference on Electronic Government and the Information Systems Perspective*, pages 131–142. Springer.
- de Sousa, R. N. and Prata, D. N. (2019). Resumo automático de textos jurídicos usando grafos com vocabulário controlado e algoritmo k-means com words embedding. *REVISTA ESMAT*, 11(18):65–80.
- de Souza Tavares, R., Lisboa, L. L. A., Zeferino, A. F., Duarte, R. P., and Valverde, B. O. (2020). Judicialização da pandemia de covid-19 no brasil: Estudo de caso sobre a aplicação do método de modelagem de tópicos para agrupamento de documentos jurídicos. *Revista Culturas Jurídicas*, 7(18).
- Divino, S. B. S. and Magalhães, R. A. (2020). Inteligência artificial e direito empresarial: Mecanismos de governança digital para implementação e confiabilidade. *Revista dos Tribunais— vol*, 1021(2020):191–212.
- Dutra, L. A., Stigert, C. C. S., da Silva Pacheco, L. A., et al. (2021). Instrução assistida de pareceres sobre processos judiciais: Assistente conjur. *Revista do TCU*, 1(148):95–103.
- Goldberg, Y. (2016). A primer on neural network models for natural language processing. *Journal of Artificial Intelligence Research*, 57:345–420.
- Josey, A., Lankhorst, M., Band, I., Jonkers, H., and Quartel, D. (2016). An introduction to the archimate® 3.0 specification. *White Paper from The Open Group*.
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *nature*, 521(7553):436–444.
- Martim, H. d., Lima, J. A. d. O., and Araujo, L. C. (2018). Base de normas jurídicas brasileiras: uma iniciativa de open government data. *Perspectivas em Ciência da Informação*, 23:133–149.
- Mehr, H., Ash, H., and Fellow, D. (2017). Artificial intelligence for citizen services and government. *Ash Cent. Democr. Gov. Innov. Harvard Kennedy Sch.*, no. August, pages 1–12.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26.

- Pedruzzi, P. (2020). Reflexões sobre o impacto da inteligência artificial na prática jurídica. *Anais dos Congressos Estaduais de Magistrados-RS*, 1(1).
- Rao, D. and McMahan, B. (2019). *Natural language processing with PyTorch: build intelligent language applications using deep learning*. "O'Reilly Media, Inc."
- Schwaber, K. and Sutherland, J. (2011). The scrum guide. *Scrum Alliance*, 21(1).
- Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A. Y., and Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 conference on empirical methods in natural language processing*, pages 1631–1642.
- Sousa, A. W. and Del Fabro, M. D. (2019). Iudicium textum dataset uma base de textos jurídicos para nlp. In *XXXIV Simpósio Brasileiro de Banco de Dados: Dataset Showcase Workshop, SBBD*.
- Souza, F., Nogueira, R., and Lotufo, R. (2020). BERTimbau: pretrained BERT models for Brazilian Portuguese. In *9th Brazilian Conference on Intelligent Systems, BRACIS, Rio Grande do Sul, Brazil, October 20-23 (to appear)*.
- Zhang, Y. and Wallace, B. (2015). A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification. *arXiv preprint arXiv:1510.03820*.