



Instrução Assistida de Pareceres sobre Processos Judiciais: assistente Conjur

Assisted Instruction of Opinions on Legal Proceedings:
Conjur Assistant

Luís André Dutra e Silva

Servidor do TCU. Bacharel em Ciência da Computação. IEEE Professional Software Engineering Master. Stanford Certified Project Manager.

Caio Cesar Sobreira Stigert

Servidor do TCU. Ex-servidor do MPU/PGR e da DPRJ. Bacharel em Direito e Relações Internacionais. Atua como Especialista-sênior no desenvolvimento do Assistente Conjur.

Leonardo Augusto da Silva Pacheco

Graduado em Ciência da Computação e em Matemática pela Universidade de Brasília, com especializações em Administração de Sistemas de Informação (FGV), Governança de TIC (UnB/ISC) e Análise de dados para o Controle (ISC). Servidor do Tribunal de Contas da União há 13 anos, Atuação em banco de dados e em desenvolvimento de software, atualmente com maior foco na liderança de projetos em recuperação da informação e em inteligência artificial.

RESUMO

A solução de instrução assistida de pareceres sobre processos judiciais da Conjur é apresentada: seus objetivos, seu impacto na mudança de paradigmas, as oportunidades de utilização de técnicas de Inteligência Artificial, a criação de um banco de conhecimento capaz de substituir os modelos de documentos sobre temas repetitivos, as sugestões de teses jurídicas pertinentes a cada processo, o preenchimento automático de minutas de pareceres e os resultados obtidos. Além disso, são enumeradas novas possibilidades de aplicação da solução em outras áreas do Tribunal.

Palavras-chave: Instrução assistida. Pareceres jurídicos. Inteligência artificial. Banco de conhecimento. Grafos. Probabilidades. Inferência. Aprendizado.



ABSTRACT

Conjur's assisted instruction solution for legal proceedings caused an impact on changing paradigms due to the opportunities to use Artificial Intelligence techniques, the creation of a knowledge bank capable of replacing document templates on repetitive topics, and suggestions of legal thesis pertinent to each process.

Keywords: Assisted instruction. Legal opinions. Artificial intelligence. Knowledge database. Graphs. Probabilities. Inference. Learning.

1. INTRODUÇÃO

O objetivo inicial do projeto foi a substituição do método tradicional de criação de pareceres jurídicos sobre temas repetitivos por algoritmos avançados de representação, inferência e aprendizado automáticos, capazes de acelerar a realização do trabalho diário dos servidores do TCU na instrução desses pareceres.

O escopo do projeto abrange cerca de 90% dos pareceres elaborados pela Conjur. Portanto, há um ganho significativo de produtividade com o uso do assistente.

A tecnologia utilizada compreende o estado da arte em Inteligência Artificial para a realização de modelagem de tópicos e similaridade de documentos de petição inicial nos processos do TCU. As técnicas utilizadas foram Processamento de Linguagem Natural (NLP), *Machine Learning* e *Deep Learning*.

Processamento de Linguagem Natural é uma tecnologia aplicada ao reconhecimento de padrões existentes em textos de documentos.

Machine learning é um conjunto de métodos que são capazes de resolver problemas sem a programação explícita de um sistema, pois todas as inferências são baseadas nos dados de treinamento dos modelos.

Deep learning é uma das aplicações de *machine learning* na qual são utilizados modelos, denominados redes neurais profundas, que trabalham com grande quantidade de dados de treinamento para realizar atividades inteligentes próximas às de um ser humano.

Análise exploratória de dados consiste em descrever de forma gráfica e textual os dados com que se pretende trabalhar em um algoritmo preditivo.

Como poderá ser visto mais adiante, o objetivo inicial de utilizar apenas temas repetitivos foi superado pela possibilidade de a ferramenta auxiliar o trabalho de instrução, na maioria dos pareceres jurídicos elaborados pela Conjur.



2. METODOLOGIA

2.1 HISTÓRICO DA INSTRUÇÃO DE PARECERES

Figura 1 - pesquisa em documentos Word



Fontes: STI, Freepik

Conforme ilustrado pela figura 1, os servidores do TCU pesquisavam o repositório de teses, ou os arquivos da Conjur, de forma sequencial e inseriam manualmente os trechos mais relevantes no parecer final.

Apenas modelos de documentos Word serviam como repositório das teses a serem utilizadas nos pareceres finais.

2.2 OPORTUNIDADE DE MELHORIA

Foi identificado o uso potencial de Inteligência Artificial para representar o conhecimento jurídico, realizar inferências e aprender automaticamente com base no resultado dos tópicos escolhidos para compor pareceres. A figura 2 representa o avanço metodológico de pesquisa em banco de conhecimento.

Figura 2 – pesquisa em banco de conhecimento



Fontes: STI, Freepik

Segundo o estudo inicial, a automatização não prescindiria do envolvimento ativo do servidor do TCU na confecção dos pareceres; portanto, apenas parte do processo de trabalho seria automatizado.

2.3 CRIAÇÃO DO ASSISTENTE

O processo de desenvolvimento do Assistente Conjur foi conduzido a partir da utilização de novas técnicas tanto na área jurídica quanto na área de TI. A solução utiliza-se das mais modernas tecnologias disponíveis no mercado de *lawtechs*, adaptando à realidade do serviço público as poderosas ferramentas de **análise e processamento de dados, jurimetria e peticionamento assistido**. Dessa forma, está alinhada ao propósito da Estratégia Digital do TCU, que busca redesenhar a forma de atuação do Tribunal no contexto digital para alavancar o impacto das ações de controle para a sociedade.

O conhecimento jurídico necessário para a criação de pareceres sobre processos judiciais pode ser representado fielmente por um banco de conhecimento que capture as peculiaridades de cada tese.

A natureza dinâmica desse banco é determinada pela constante atualização das teses jurídicas e, conseqüentemente, pela atualização do modelo de predição da solução.

O banco de conhecimento também fornece a possibilidade de associar perguntas à exibição ou não de textos específicos, bem como a associação de variáveis a metadados dos pareceres.

Cada tese armazenada nesse banco serve como entrada em um modelo de tópicos que captura a essência do texto e realiza sugestões de teses com base em semelhança semântica com a petição inicial de cada processo.

Outro algoritmo de sugestão de teses é o agrupamento por afinidade, que leva em consideração o ponto de vista de cada usuário em relação a confirmar ou excluir determinada tese do conjunto apresentado. Dessa forma, o aprendizado ocorre em tempo real para esse caso específico.



A classificação de tópicos de forma supervisionada soma-se aos demais algoritmos como mais uma opção de seleção dos tópicos mais relevantes para cada processo.

Esses algoritmos de aprendizado são executados previamente em lote, durante a noite, com a finalidade de fornecer agilidade aos servidores do TCU durante a elaboração de pareceres. Durante esse processo, as escolhas realizadas durante o dia são consolidadas pelo treinamento noturno.

Grande parte dos dados usados no preenchimento dos pareceres da Conjur estão presentes nos sistemas e-TCU e Sistema de Gestão de Demandas da Conjur (SGDC).

Com isso, a geração de minutas de pareceres formatados de acordo com o padrão de cada tipo de processo é realizada de forma automática, evitando ao máximo grandes mudanças no texto final.

Para determinar a exibição de determinados trechos do parecer, são formuladas perguntas no momento em que o servidor escolhe um ou mais tópicos de uma tese.

Além disso, certos metadados podem ser preenchidos automaticamente, por meio de inferência exata, ou manualmente.

3. RESULTADOS

O Assistente Conjur trouxe, sob vários aspectos, melhorias significativas no processo de trabalho. Em primeiro lugar, a padronização das teses e dos modelos proporcionou maior confiabilidade e uniformidade nos pareceres exarados pela Conjur, evitando divergências entre o trabalho dos servidores. A atualização dos modelos diretamente no Banco de Conhecimento permite que o servidor sempre tenha acesso ao modelo mais atual, sem a necessidade de se preocupar se está usando o texto correto ou a jurisprudência mais atualizada para determinada matéria.

O formato do Banco de Conhecimento também permite que sejam feitas atualizações de forma mais célere, incorporando de forma mais fluida, na rotina de trabalho, os novos temas e as alterações sobre temas existentes. Tudo isso traz mais segurança e agilidade para o trabalho dos servidores.

Além disso, o ganho de tempo na confecção dos pareceres é significativo. O preenchimento automático de metadados e dados específicos do processo bem como o preenchimento automático das teses, com o auxílio das perguntas e das variáveis, e a formatação automática dentro do *template* utilizado pela Conjur viabilizam a produção de instruções em muito menos tempo se comparado ao antigo processo de trabalho, bastando ao servidor, muitas vezes, fazer apenas uma revisão da minuta que o Assistente gerou.

Em termos quantitativos, conforme demonstrado pela tabela 1, os ganhos de eficiência variam em razão da complexidade do processo, da qualidade das peças encaminhadas à Conjur e da



recorrência da matéria. De plano, verificou-se uma economia de tempo entre 30 minutos e 1 hora apenas com a automatização da formatação e o preenchimento de metadados.

Além disso, para processos repetitivos, com teses e argumentos bem delineados e cuja matéria a IA habituou-se a tratar em razão do volume de casos processados, a razão produtividade/tempo recebe um incremento de 70% a 80%, reduzindo o trabalho do servidor à mera revisão da minuta e de ajustes pontuais. Nesses casos, uma minuta de parecer ou instrução que demoraria de 6 a 8 horas para ser produzida – considerando-se o tempo de análise, discussão e redação – hoje pode ser confeccionada em 1 hora.

Por outro lado, para processo mais complexos – *leading cases*, processos que demandem análise de muitos documentos e anexos, processos com temas sensíveis ou estratégicos ou processos com muitas particularidades concretas – o incremento de produtividade ficou entre 20% e 50%, a depender dos já mencionados fatores objetivos influentes no funcionamento do Assistente.

Mantendo a variável externa “qualidade textual das peças analisadas” constante (qualidade razoável/boa), podemos resumir o incremento de produtividade no seguinte quadro:

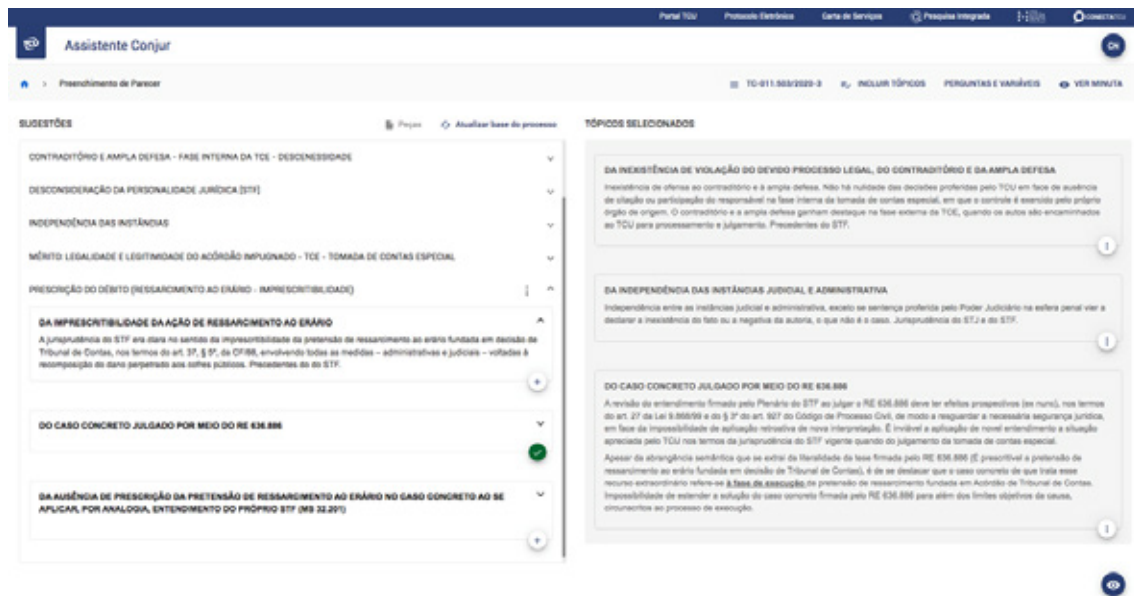
Tabela 1 – resultados obtidos com o assistente

| Tipo de processo | Incremento de produtividade | Tempo médio de produção com os métodos tradicionais | Tempo médio de produção com o Assistente |
|---|-----------------------------|---|--|
| Temas repetitivos | 80% | 6 a 8 horas | 1 a 2 horas |
| Processos de baixa complexidade | 70% | 8 a 12 horas | 2 a 4 horas |
| Temas estrategicamente sensíveis repetitivos e/ou de baixa complexidade | 50% | 2 dias | 1 dia |
| Temas estrategicamente sensíveis de alta complexidade | 30% | 3 dias | 2 dias |
| Processos que demandem análise minuciosa de documentos e aspectos fáticos | 40% | 2 dias | 1 dia a 1 dia e meio |
| Leading Cases de baixa complexidade | 30% | 2 a 3 dias | 1 a 2 dias |
| Leading Cases de alta complexidade ou estrategicamente relevantes | 20% | 4 a 5 dias | 3 a 4 dias |



Para além de tais ganhos mensuráveis, há, ainda, um enorme benefício intangível verificado a médio e longo prazo relacionado à sistematização, uniformização e gestão do banco de conhecimento, que viabiliza respostas muito mais rápidas, eficazes e assertivas aos novos paradigmas doutrinários e jurisprudenciais. Nesse escopo, o Assistente elimina a necessidade de atualização de documentos já produzidos ou modelos parametrizados, viabilizando que qualquer alteração pontual ou substancial em determinado tema ou argumento fique imediatamente disponível para todos os usuários do sistema após a atualização pelo gestor.

Figura 4 – interface de elaboração de pareceres



Fonte: STI

A figura 4 mostra a interface de elaboração de pareceres do Assistente Conjur. A localização das teses, por meio de IA, direciona o servidor para os melhores modelos sobre o assunto tratado no processo; assim como a modulação textual, a partir de perguntas e variáveis, proporciona uma personalização em níveis mais profundos, de forma que cada vez menos alterações sejam necessárias no parecer final.

Nesse contexto, é importante pontuar o ganho de produtividade que o Assistente proporciona nos casos repetitivos, visto que, muitas vezes, há um volume grande de processos com viés de análise similar, em que o texto varia pouco ou quase nada.

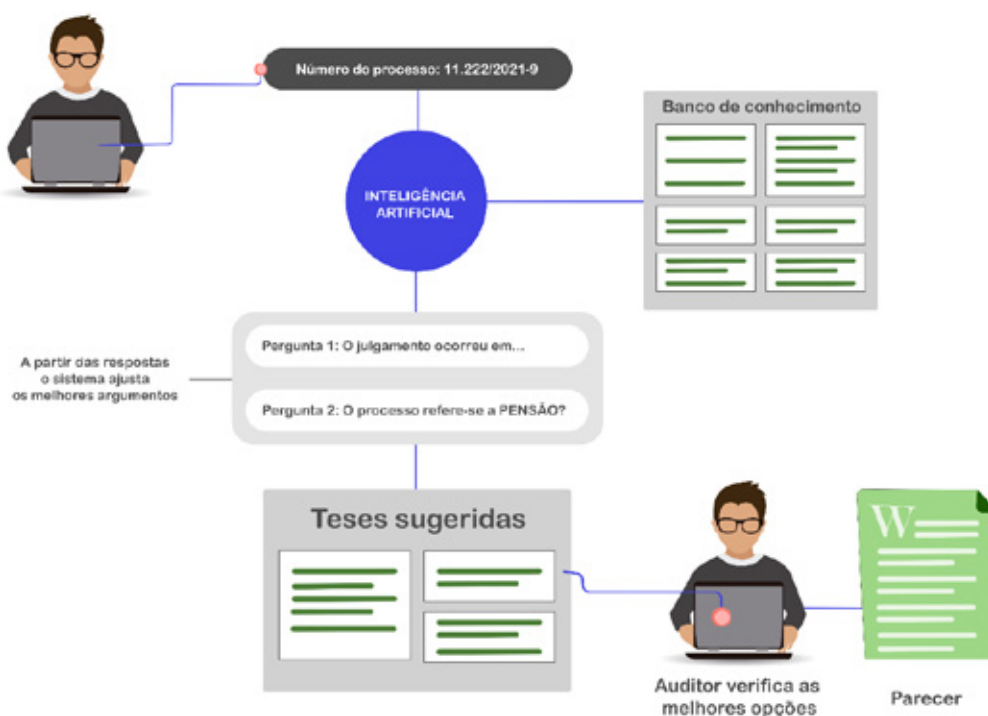
Nesses casos, o Assistente consegue produzir um número muito maior de instruções por unidade de tempo, desonerando o servidor para se ater a casos mais complexos.

Na área jurídica, destaca-se a reorganização dos modelos, adequando-os ao que se chama de “modelagem por argumentos”. A construção argumentativa sugerida pelo Assistente segue um padrão de árvores lógicas, com infinitas possibilidades de personalização.

Na área de TI, destaca-se o uso do *framework Probabilistic Graphical Models (PGM)*, ou, mais especificamente, *Markov Random Fields*, para o aprendizado da afinidade existente entre as teses sugeridas. Além desse, o uso de *word embeddings* para expansão semântica das teses e a modelagem de tópicos por meio do algoritmo LDA (*Latent Dirichlet Allocation*), bem como aprendizado supervisionado usando o algoritmo *Random Forest*. Por fim, o sistema utiliza um banco de grafos para a modelagem do Banco de Conhecimento.

Dentro desse contexto, foi possível utilizar a técnica de perguntas e variáveis – por meio da qual, a partir de respostas dadas pelo servidor do TCU, o sistema ajusta os melhores argumentos e os metadados para o caso específico – tornando as minutas produzidas cada vez mais acuradas e particulares.

Figura 5 – fluxo de elaboração de pareceres



Fontes: STI, Freepik

Além disso, a integração com o e-TCU e o SGDC e com a Pesquisa Integrada do Tribunal torna o processo de trabalho mais fluido e conectado, e extrai o máximo das funcionalidades de cada sistema para proporcionar uma experiência mais completa.

3.1 POSSIBILIDADES DE DISSEMINAÇÃO DA SOLUÇÃO PARA OUTRAS ÁREAS

O formato e as tecnologias adotadas na criação do Assistente Conjur viabilizam, com as devidas adaptações, a utilização do sistema por várias unidades do TCU. De imediato, vislumbra-se a possibilidade de ajudar na confecção de pareceres, de respostas a petições e de recursos.



Com o desenvolvimento em curso do novo módulo de Licitações e Contratos, também unidades como Segedam e Selip poderão ser beneficiadas, visto que o Assistente contará com um sistema de aprimoramento da rotina de trabalho, tratando eventuais falhas de editais e contratos antes mesmo que eles cheguem à Conjur.

Em resumo, quaisquer unidades que trabalhem com análise de documentos, processos, editais, contratos, votos, pareceres, acórdãos, instruções ou comunicações poderão ser beneficiadas com o uso das funcionalidades do Assistente.

4. CONCLUSÃO

O Assistente Conjur representa um salto de produtividade e amplia as perspectivas para o uso de ferramentas de IA em outras áreas do Tribunal e do serviço público como um todo.

Esse produto, além de reduzir o tempo despendido na elaboração de pareceres e informações, possibilita melhor gestão do conhecimento e maior padronização nas atividades da Consultoria Jurídica e contribui significativamente para o incremento da eficiência no âmbito da unidade.

REFERÊNCIAS BIBLIOGRÁFICAS

ENGELMANN, Wilson; WERNER, Deivid Augusto. Inteligência Artificial e Direito. In: Inteligência artificial e direito: ética, regulação e responsabilidade; coordenação Ana Frazão e Caitlin Mulholland. São Paulo: Thomson Reuters Brasil, 2019.

GIRARDI, Rosario. Inteligência artificial aplicada ao Direito. Clube de Autores, 1ª edição, 2015.

FERRARI, Isabela; LEITE, Rafael; RAVAGNANI, Giovani; FEIGELSON, Bruno. Justiça Digital; coordenação Isabela Ferrari. São Paulo: Thomson Reuters, Revista dos Tribunais, 1ª edição, 2020.

C. Hsu and C. Chiu, "A hybrid Latent Dirichlet Allocation approach for topic classification," 2017 IEEE International Conference on INnovations in Intelligent SysTems and Applications (INISTA), Gdymia, 2017, pp. 312-315, doi: 10.1109/INISTA.2017.8001177.

L. G. Mojica and V. Ng, "Fine-Grained Opinion Extraction with Markov Logic Networks," 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), Miami, FL, USA, 2015, pp. 271-276, doi: 10.1109/ICMLA.2015.215.

M. A. Silva-Fuentes, H. D. Calderon-Vilca, E. F. Calderon-Vilca and F. C. Cárdenas-Mariño, "Semantic Search System using Word Embeddings for query expansion," 2019 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America), Gramado, Brazil, 2019, pp. 1-6, doi: 10.1109/ISGT-LA.2019.8894992.

Os conceitos e interpretações emitidos nos trabalhos assinados são de exclusiva responsabilidade de seus autores.

