

1. IDENTIFICAÇÃO DO DOCUMENTO

Nome do Projeto:	CLASSIFICADOR DE TEMAS	Código:	
Demandante:	CSJT		

2. NOMEAÇÃO DO GERENTE DO PROJETO

Gerente do Projeto:	ANA CAROLINA PEREIRA ROCHA MARTINS
Unidade Administrativa:	SPTRI - CTPJE
Telefone:	61 3043 7929
E-Mail:	acprocha@tst.jus.br

3. HISTÓRICO DE REVISÕES

Data	Versão	Descrição	Responsável
30/05/2017	1.0	Versão inicial	Ana Carolina P. Rocha Martins

4. DOCUMENTOS DE REFERÊNCIA

Plano Estratégico do Conselho Superior da Justiça do Trabalho - 2015 -2020. Disponível em <http://www.csjt.jus.br/c/document_library/get_file?uuid=a194cab-a556e-4cf6-96db-78c0f21b1438&groupId=955023>

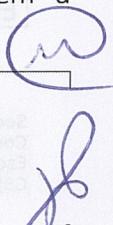
Plano Estratégico da Justiça do Trabalho - 2015 -2020. Disponível em <http://www.csjt.jus.br/c/document_library/get_file?uuid=f525e749-2197-438c-91ae-d31acf4cbdf&groupId=955023>

FELDMAN, R., & SANGER, J. (2007). **The text mining handbook: advanced approaches in analyzing unstructured data.** Cambridge university press.

CARVALHO, Maximiliano (2017). **O Princípio Da Automatização Do Processo Eletrônico Como Catalisador Da Observância Aos Precedentes Do TST.** Revista Fórum Trabalhista RTF, Nº 24. Disponível em <<http://www.editoraforum.com.br/ef/index.php/publicacoes/periodicos/listar-periodicos/revista-forum-trabalhista-rft/>>

5. JUSTIFICATIVA DO PROJETO

O crescimento do volume de processos trabalhistas e do número de documentos dentro do PJe tem tornado análise manual desta quantidade de dados digitais um limitador para a agilidade de tramitação dos processos. Assim, se faz necessário que o PJe KZ adquira funcionalidades baseadas em inteligência artificial que possam auxiliar o homem a



trabalhar com a ferramenta de forma mais eficiente, maximizando-se “a produção com menor consumo de energia, de modo a entregar quantitativa e qualitativamente mais rápido a prestação jurisdicional, ao mesmo tempo em que se melhora a qualidade de vida dos usuários internos e externos do processo eletrônico” (CARVALHO, 2017), poupando ao homem o trabalho que uma máquina pode fazer.

À vista disto, este projeto propõe uma automação de máquina que, a partir da leitura dos documentos de decisões do segundo grau de processos com tema identificado, possa aprender o vocabulário utilizado para abordar cada tema, e a partir disto, classifique de forma automática novos processos como tema ainda não identificado em segundo e terceiro grau. Foram escolhidos os temas, ao invés de assuntos, por serem uma informação mais detalhada e útil a justiça trabalhista.

Uma vez implementado este projeto, uma série de funcionalidades poderão ser construídas no PJe KZ, tais como triagem por temas, busca textual em peças do processo, busca textual semântica em peças do processo, sugestão de análise de precedentes no momento da redação de decisões, criação de avisos para a identificação de temas de processos repetitivos, sugestão de análise de processos similares no momento da conclusão de uma decisão, entre outros.

6. OBJETIVO DO PROJETO

Construir um aprendizado de máquina que seja capaz de reconhecer informações relevantes nas peças processuais e, a partir destas informações, identificar os temas ali abordados de forma automática.

7. ALINHAMENTO ESTRATÉGICO

Origem (PEI/PDTIC/PET IC)	Objetivo	Indicador	Impacto no indicador
Plano Estratégico 2015- 2020 do CSJT	Contribuir para a melhoria da prestação jurisdicional na Justiça do Trabalho de 1º e 2º graus	Índice de Satisfação Interna com o Sistema do Processo Judicial Eletrônico (ISIPje)	Influenciar positivamente no número de questionários que aprovam o sistema PJe.
Plano Estratégico	Contribuir para a melhoria da prestação jurisdicional	Índice de Satisfação Externa com o	Influenciar positivamente



2015-2020 do CSJT	na Justiça do Trabalho de 1o e 2o graus	Sistema do Processo Judicial Eletrônico (ISEPJe)	número de questionários que aprovam o sistema PJe, uma vez que haverá maior celeridade com a implementação do projeto.
Plano Estratégico da Justiça do Trabalho 2015 -2020	Assegurar a celeridade e a produtividade na prestação jurisdicional	Tempo Médio de Duração do Processo - 2a Instância (TMDP2)	Diminuir o tempo de tramitação do processo, o que impacta diretamente no tempo médio de duração do processo.
Plano Estratégico da Justiça do Trabalho 2015 -2020	Assegurar a celeridade e a produtividade na prestação jurisdicional	Índice de Processos Julgados (IPJ)*	Ao trazer agilidade ao processo de julgamento, poderá aumentar o número de processos julgados

8. DESCRIÇÃO DO ESCOPO

- Serviço de extração dos documentos do PJe em formato de texto, para que possa ser feita a indexação destes;
- Pré-processamento dos textos extraídos (remoção de tags html por exemplo);
- Indexação dos textos do PJe;
- Criação de uma estrutura de dados que contenha os temas uniformizados pelo Projeto Unificação dos Temas Processuais Para Toda a Justiça do Trabalho¹. Caso não haja uma priorização do projeto Projeto Unificação dos Temas Processuais Para Toda a Justiça do Trabalho, os temas serão buscados do Banco Nacional de Jurisprudência Trabalhista (BANJUR) e dos temas de Recursos Repetitivos;
- Criação de um modelo de mineração de textos capaz de classificar o tema dos processos que já se encontram classificados (aprendizado);
- Aplicação do modelo de mineração de textos para a classificação dos processos das bases dos tribunais que aderirem à solução (aplicação do aprendizado);

¹ O Projeto Unificação dos Temas Processuais Para Toda a Justiça do Trabalho, ganhador do concurso “Projetos e Oportunidades” do Tribunal Superior do Trabalho (TST) de 2017, se propõe a estruturar e uniformizar os temas que são utilizados na classificação de processos nos órgãos judicantes deste órgão.

- Criação de uma tela para alteração do tema classificado, para o caso de haver necessidade de retificação ou complementação.
- Integração com o PJe KZ;
- Guia de instalação;
- Criação do manual referente à funcionalidade.

8.1. Não escopo do projeto

- Criação de novas funcionalidades baseadas na indexação dos textos e no aprendizado construído;
- Criação de um *thesaurus* jurídico (dicionário de conceitos negociais);
- Unificação dos temas processuais da justiça do trabalho;
- Criação de novos temas;

8.2. Premissas Iniciais

- Assinatura da Carta de Compromisso de Liberação, até o dia 1º de junho de 2017 (não prorrogável), para que seja oficializada a permissão para a servidora Ana Carolina Pereira Rocha Martins se ausentar às sextas feiras de sua unidade de trabalho, durante o período de 24 meses, para que possa comparecer às aulas ministradas na UNB e empreender demais estudos do mestrado necessários à conclusão do projeto;
- Permissão para que, a partir do primeiro semestre de 2018, a servidora Ana Carolina Pereira Rocha Martins atue em demandas relacionadas a este projeto em conjunto com as demais demandas de sua atribuição ordinária;
- Aprovação da servidora Ana Carolina Pereira Rocha Martins em todas as matérias do programa de mestrado, com tolerância a apenas uma reprovada;
- Apoio de uma pessoa ou grupo de negócio detentor do saber jurídico inerente ao projeto para apoio do saber negocial e validação da solução. Não há necessidade inicial de dedicação integral ao projeto;
- Apoio da SISUP para auxílio da disponibilização da solução. Não há necessidade inicial de dedicação integral ao projeto;
- Apoio da Seção de Gestão do Produto - SGPROD - para auxílio da validação da solução. Não há necessidade inicial de dedicação integral ao projeto;

- Apoio de uma pessoa da Seção de Métodos e Padrões (SMPAD) para consolidação da arquitetura e apoio no desenvolvimento. Não há necessidade inicial de dedicação integral ao projeto;
- Apoio de uma pessoa da Seção de Análise e Projetos para Tribunais (SPTRI) no desenvolvimento o projeto. Não há necessidade inicial de dedicação integral ao projeto;
- Priorização do projeto por parte da gestão do CSJT dos biênios de 2016-2017 e 2018-2019.
- Acesso às bases bugfix de todos os TRTs, incluindo as bases binárias;
- Acesso aos dados do Projeto Unificação dos Temas Processuais Para Toda a Justiça do Trabalho;
- Caso o Projeto Unificação dos Temas Processuais Para Toda a Justiça do Trabalho não seja implementado, será preciso ter acesso à base do BANJUR e aos temas de recursos repetitivos do TST e dos regionais.

8.3. Restrições Iniciais

- Considerando a elevada influência dos dados com os algoritmos de *text mining*² no desempenho dos algoritmos, segundo FELDMAN & SANGER (2007) , a performance desses métodos só pode ser medida em caráter experimental, se comportando de forma diferente para cada cenário de aplicação. Neste caso, embora não se possa determinar o percentual do índice de acerto, é sabido que uma margem de erro é esperada neste tipo de técnica. Assim, fica aqui o registro que esta solução não terá 100% de eficácia.



2 Área da inteligência artificial que se propõe a resolver a crise da sobrecarga de informação extraíndo as informações relevantes de textos ao combinar técnicas de mineração de dados, aprendizado de máquina, processamento de linguagem natural, recuperação da informação e gerenciamento do conhecimento (FELDMAN & SANGER, 2007)



9. CRONOGRAMA DE MARCOS

Como se trata de um projeto de dissertação de mestrado profissional, o prazo estipulado pelo Programa de Pós Graduação em Ciência Aplicada da UNB é de 2 anos a partir do ingresso como aluno regular, podendo se estender por mais 6 meses, se necessário. Este então é o balizador do prazo necessário para a implementação solução. As datas aqui apresentadas são apenas uma estimativa do tempo de desenvolvimento da solução, podendo ser alteradas à medida que forem sendo desenvolvidas. À pedido da Coordenação Nacional Executiva do PJe(CNE), a implementação do projeto dentro da Coordenação Técnica do PJe (CTPJE) terá início em 2018.

Marcos	Previsão
Criação de serviço para disponibilização das peças processuais em formato de texto para a ferramenta de indexação	Abril/2018
Instalação de uso da ferramenta de indexação no ambiente de desenvolvimento (Ex: Solr, Elasticsearch)	Junho/2018
Qualificação da dissertação junto à UNB	Junho/2018
Separação de processos que serão usados para o treinamento	Julho/2018
Implementação de algoritmos de mineração de textos para a descoberta do modelo de classificação	Fevereiro/2019
Aplicação dos conhecimentos construídos em novas peças processuais dentro do PJe KZ	Março/2019
Homologação da solução no ambiente de desenvolvimento/homologação	Maio/2019
Realimentação da base de conhecimento com as novas decisões julgadas	Junho/2019
Implantação em produção em um Tribunal Regional do Trabalho para testes	Junho/2019
Implantação nos demais regionais	Agosto/2019

10. RISCOS PREVIAMENTE IDENTIFICADOS

Probabilidade		Impacto (Efeito que o risco exerce sobre o projeto)		Grau de Risco	
Índice	Probabilidade de Ocorrência	Índice	Impacto	Índice	Descrição
1	Improvável	1	Muito baixo	1 a 2	Muito Baixo - Impacto mínimo no projeto
2	Pouco provável	2	Baixo	3 a 5	Baixo - Impacto no projeto
3	Provável	3	Médio	6 a 10	Médio - Impacto no projeto
4	Muito provável	4	Alto	12 a 16	Alto - Impacto no projeto
5	Quase Certo	5	Muito alto	20 a 25	Muito Alto - Comprometimento no projeto

Tabela de identificação e plano de resposta aos riscos:

Nº	Risco Encontrado	Probabilidade	Impacto	Grau do Risco	Ação Contingência de	Responsável pela ação de contingência
1	Impossibilidade de assinar a Carta de Compromisso de Liberação até 1º de Junho de 2017	3	5	15	Adiamento no ingresso no mestrado por um ano, com possibilidade de atraso no cronograma devido à falta de conhecimento necessário para a implementação do projeto.	-
2	Impossibilidade de permitir que a servidora Ana Carolina atue em demandas relacionadas ao projeto em conjunto com as demais demandas	2	4	8	Não há. Algumas datas provavelmente serão postergadas.	CNE e CTPje

3	Reprovação da servidora Ana Carolina em mais de uma matéria no programa de mestrado da UNB	2	5	10	Continuação do projeto sem parceria com a UNB com possibilidade de atraso no cronograma devido à falta de conhecimento necessário para a implementação do projeto.	Ana Carolina Pereira Rocha Martins
4	Impossibilidade de se ter uma pessoa disponível que possa responder pelo conhecimento negocial	2	5	10	Construção do conhecimento de forma autodidata e buscando conhecimento com pessoas da área jurídica.	Ana Carolina Pereira Rocha Martins
5	Impossibilidade da SISUP atuar no projeto	2	5	10	Não há. A solução não poderá ser implantada em produção.	CNE e CTPJe
6	Impossibilidade da SGPROD atuar no projeto	2	4	8	A validação da solução terá que acontecer pela pessoa de apoio negocial	CNE e CTPJe
7	Impossibilidade de que a SMPAD disponibilize alguém para auxiliar na proposta de arquitetura do projeto	2	3	6	Uma arquitetura será proposta, mas poderá não ser a mais adequada ao projeto PJe.	Ana Carolina Pereira Rocha Martins
8	Impossibilidade de que a SPTRI disponibilize alguém para auxiliar no desenvolvimento do projeto	2	4	8	Disponibilizar alguém da Seção de Análise e Projeto para Varas (SPVAR) ou SMPAD. Não sendo possível, poderá haver atraso no cronograma.	CNE e CTPJe

9	Troca da gestão e despriorização do projeto	4	5	20	Não há. É necessário que este projeto esteja priorizado pela CNE.	Presidente do CSJT;CNE e CTPJe
10	Impossibilidade de acesso às bases bugfix dos TRTs	2	5	10	Não há. O projeto necessita disto.	CNE e CTPJe
11	Impossibilidade de acesso aos dados do Projeto Unificação dos Temas Processuais Para Toda a Justiça do Trabalho	2	4	8	Concessão de acesso aos temas e processos presentes no BANJUR e aos temas e processos de recursos de revista repetitivos e demandas repetitivas.	CNE e CTPJe
12	Projeto Unificação dos Temas Processuais Para Toda a Justiça do Trabalho não concluído a tempo	3	4	12	Concessão de acesso aos temas e processos presentes no BANJUR e aos temas e processos de recursos de revista repetitivos e demandas repetitivas.	CNE e CTPJe
13	Impossibilidade de acesso aos dados do BANJUR e dos temas de processos repetitivos	2	5	10	Trocando a abordagem. Ao invés de se fazer um aprendizado de máquina para a classificação posterior de novas peças, faz-se o reconhecimento de citações de súmulas, dispositivos (entre outras entidades relevantes) para que se faça o agrupamento de processos similares	Ana Carolina Pereira Rocha Martins

14	Não confiabilidade da classificação de temas/assuntos existente atualmente (BANJUR, recursos repetitivos)	3	5	15	Trocara abordagem. Ao invés de se fazer um aprendizado de máquina para a classificação posterior de novas peças, faz-se o reconhecimento de citações de súmulas, dispositivos (entre outras entidades relevantes) para que se faça o agrupamento de processos similares	Ana Carolina Pereira Rocha Martins
15	Baixa quantidade de processos categorizados corretamente, não sendo suficiente para construir a inteligência de classificação	3	4	12	Trocara abordagem. Ao invés de se fazer um aprendizado de máquina para a classificação posterior de novas peças, faz-se o reconhecimento de citações de súmulas, dispositivos (entre outras entidades relevantes) para que se faça o agrupamento de processos similares	Ana Carolina Pereira Rocha Martins
16	Impossibilidade de disponibilizar máquina com processamento adequado	3	4	12	Construir a solução de forma distribuída (clusters)	Ana Carolina ,SISUP e equipe de infraestrutura do regional
17	Na impossibilidade de disponibilizar máquina com processamento adequado, impossibilidade de construir a aplicação de forma distribuída	3	5	15	Não há. É necessário que se tenha capacidade computacional disponível.	-

11. PRINCIPAIS ENVOLVIDOS NO PROJETO

Nome	Área	Papel	Telefone(s)	E-mail
Dr. Maximiliano Pereira de Carvalho	Coordenação Nacional Executiva do PJe (CNE)/ CSJT	Patrocinador	061 3043 7384	maximiliano.carvalho@tst.jus.br
Herbert Parente	CTPJE/ CSJT	Coordenador Técnico	061 3043 7711	herbert.parente@tst.jus.br
Ana Carolina Pereira Rocha Martins	SPTRI/ CTPJE/ CSJT	Desenvolvedora, líder técnica do projeto	061 3043 2979 061 99927 5063	acprocha@gmail.com
Glauber Moreira Rocha	SPTRI/ CTPJE/ CSJT	Apoio no desenvolvimento	061 3043 7824	glauber.rocha@tst.jus.br
Christiano Guimaraes de Carvalho	SISUP/ CTPJE/ CSJT	Atuar na disponibilização da solução	061 3043 7927	christiano.carvalho@tst.jus.br
Victor Lopes Dias de Araújo	SGPROD /CTPJE/ CSJT	Validação da demanda	061 3043 7941	victor.araujo@tst.jus.br
Daniel Souto Rocha	SMPAD/ CTPJE/ CSJT	Apoio na definição da arquitetura	061 3043 7582	daniel.rocha@tst.jus.br
A definir	Grupo de Negócios	Apoio negocial	A definir	A definir

12. EQUIPE TÉCNICA INICIAL DO PROJETO

Nome	Papel	Telefone(s)	E-mail
Ana Carolina Pereira Rocha Martins	Analista/Desenvolvedora	61 3043 7929	acprocha@tst.jus.br

13. APROVAÇÃO

Data	Nome	Departamento / Unidade	Assinatura
09/06/17	MAXIMILIANO CARVALHO	PRESIDENCIA CSJT	M. Carvalho
02/06/17	FABIANO COVATO DE SOUZA	PRESIDENCIA CSJT	F. Covato

Todas as páginas deverão ser rubricadas