
BOLSA DE INICIAÇÃO CIENTÍFICA – PROCESSO SELETIVO

O Escritório de Apoio à Pesquisa comunica a abertura de processo seletivo para indicação de um bolsista na modalidade iniciação científica para atuar sob orientação do Prof. Dr. Ildeberto Aparecido Rodello no âmbito do projeto “Modelagem e implementação de uma base de dados aberta para fins de análise de decisões judiciais” (ANEXO I), vinculado ao projeto principal “Mediação e Conciliação avaliadas empiricamente: jurimetria para proposição de ações eficientes”, em parceria com o Conselho Nacional de Justiça.

1. REQUISITOS

Vaga exclusiva para alunos regularmente matriculados nos cursos de graduação da USP, cumprindo os seguintes requisitos:

- a) Requisitos: Cursando Ciência da Computação ou cursos afins (a partir do terceiro ano/sexto semestre), com conhecimentos avançados em sistema de gerenciamento de banco de dados NoSQL, mais especificamente MongoDB e conhecimentos básicos em lógica de programação.
- b) Valor Mensal da Bolsa: R\$840,00
- c) Vigência da Bolsa: 12 meses a partir da contratação

2. ATIVIDADES

Veja ANEXO I

3. INSCRIÇÃO

As inscrições serão presenciais, efetuadas por meio da entrega dos documentos abaixo relacionados no Escritório de Apoio à Pesquisa da FEARP (Bloco A, sala 40, fones (16) 3315-0675, 3315-4961 e 3315-9083) no período de 23 a 31/5/2019, diariamente, das 8h às 16h:

- a) Carta de motivação indicando como o candidato atende aos requisitos necessários;
- b) Histórico escolar de graduação;

4. SELEÇÃO

A seleção será realizada pela análise dos documentos solicitados no item 3 e eventual entrevista presencial conduzida pelo coordenador do projeto.

5. RESULTADO

O resultado será divulgado até às 18h do dia 3/6/2019 na página eletrônica da Pesquisa no sítio da FEARP (<https://www.fearp.usp.br/pt-br/pesquisa>).

ANEXO 1

Identificação do Projeto

Título: Modelagem e implementação de uma base de dados aberta para fins de análise de decisões judiciais

Proponente: Prof. Dr. Ildeberto Aparecido Rodello

Palavras-chave: Data science, Banco de dados NoSQL, Jurimetria, Tomada de decisão

1. Contextualização

Atualmente, a informação tem sido reconhecida como um ativo estratégico para qualquer organização e área de conhecimento. A gestão e a liderança eficazes envolvem, principalmente, a solução criativa de problemas, motivando os funcionários e garantindo que a organização, seja ela pública ou privada, atinja seus objetivos e metas.

Os dados obtidos de fontes internas e externas se constituem nesse ativo estratégico e são um desafio particularmente grande para os modernos sistemas de Tecnologia da Informação (TI), porque tornam esses sistemas muito lentos e não escaláveis o suficiente (LAUDON; LAUDON, 2015).

No que concerne a área de Direito, o acesso à informação é a ferramenta mais fundamental para garantir acesso qualificado à justiça. Apesar da previsão constitucional existente desde 1988 e que determina o acesso à informação como um direito individual dos cidadãos, o regramento infraconstitucional começou a regulamentar essas questões tardiamente.

Por um lado, aproveitando o poder dos recursos computacionais, os advogados podem prever com mais precisão como os eventos se desenrolam nos litígios. Assim, os advogados que adotarem a tomada de decisão baseada em dados terão uma clara vantagem sobre seus colegas que ainda se apegam a seus instrumentos desatualizados. (LETTIERI et al., 2018)

Por outro lado, o Conselho Nacional de Justiça (CNJ), os Tribunais e os magistrados tem a possibilidade, com a utilização massiva de recursos e técnicas de TI, de analisar cenários e traçar estratégias de ação a fim de minimizar, ou mesmo solucionar, notórios problemas como, por exemplo, a demora na prestação jurisdicional e a consequente ineficiência do judiciário.

No cenário das políticas judiciárias, a evolução qualitativa e quantitativa dos serviços oferecidos à sociedade envolve a identificação das principais dificuldades e oportunidades existentes em relação à formulação, à implementação e ao controle das atividades de prestação jurisdicional na busca de concretização de direitos e garantias fundamentais (CNJ, [s.d.]).

Neste sentido, o Departamento de Pesquisas Judiciárias do Conselho Nacional de Justiça (DPJ/CNJ) realiza estudos e promove a produção de pesquisas com o objetivo de promover melhorias na prestação jurisdicional.

O presente projeto está inserido em um projeto de pesquisa em vigência apoiado pelo DPJ/CNJ, cujo proponente é integrante do grupo de pesquisadores. O projeto "**Mediação e Conciliação avaliadas empiricamente: jurimetria para proposição de ações eficientes**" se refere ao campo temático "Mediação e Conciliação: Obstáculos e Incentivos na Justiça Cível", contemplado no Edital de convocação pública e de seleção convocação n. 02/2017 (EDITAL 02/2017, [s.d.]).

2. Mediação e Conciliação: Obstáculos e Incentivos na Justiça Cível

A demora na prestação jurisdicional e a ineficiência do judiciário são problemas amplamente conhecidos que não colaboram para a solução de conflitos sociais. A área jurídica se mostra invariavelmente restrita ao desenvolvimento de teorias jurídicas e processuais que, sem dúvida, são necessárias para o incremento do campo do direito, porém pouco auxiliam para a desejada entrega da pacificação social. Assim, o projeto "**Mediação e Conciliação avaliadas empiricamente: jurimetria para proposição de ações eficientes**" é uma pesquisa empírica que estudará formas de aperfeiçoar a entrega da prestação jurisdicional.

Trata-se de, inicialmente, coletar todos os dados referentes aos processos em que houve conciliação ou mediação nos estados desejados (CE, SP, PR, RJ e PI). A tarefa não se demonstra simples e configura empreendimento necessário para um trabalho incontestável. Não se trata de definir uma amostragem e avaliar alguns processos. Trata-se de montar um banco de dados com todos os

Escritório de Apoio à Pesquisa

processos existentes - seja por meio do acesso a dados públicos dos tribunais ou mesmo por acesso às informações fornecidas pelo próprio Tribunal. Com base nesse banco de dados, serão feitas inicialmente análises quantitativas sobre os processos, o que fornecerá informações extremamente relevantes para a tomada de decisões e propostas de ação, como o tempo médio de duração das demandas e os valores negociados.

É apenas com base nesses dados coletados que é possível chegar à conclusão de que tipo de ação é mais eficiente e eficaz para a melhoria dos processos internos que levem a decisões com melhores resultados (mais céleres e que alcancem, de fato, a pacificação social). Assim, após o estudo do que empiricamente tem acontecido nos Tribunais, serão determinadas e propostas as ações relevantes tendentes a aperfeiçoar a dinâmica das conciliações nos tribunais do país.

2.1. Objetivo geral

Investigar o fluxo e a forma de gestão dos processos da Justiça Estadual de primeiro grau (comarcas e varas, do interior e da capital), com relação à existência de conciliação e/ou mediação.

2.2 Objetivos específicos

- Identificar e analisar os fatores relevantes para o sucesso de uma conciliação e/ou mediação;
- Compreender o fluxo de trabalho (processos de gestão) de um processo de conciliação e/ou mediação;
- Propor ações para a organização e uniformização dos serviços de conciliação e mediação;
- Propor ações educativas para aplicação à prestação jurisdicional.

2.3 Perguntas de pesquisa

Tendo em vista o objeto e os objetivos da pesquisa ora proposta, colocam-se as seguintes perguntas de pesquisa:

- 1) Qual o fluxo e a forma de gestão dos processos da Justiça Estadual de primeiro grau de conciliação e/ou mediação?
- 2) Por que algumas unidades judiciárias conciliam mais do que outras?
- 3) Há diferenças relevantes relacionadas com a quantidade de processos solucionados via conciliação e mediação (em relação ao total de processos da comarca)?

2.4 Hipóteses de pesquisa

A partir das perguntas de pesquisa acima apontadas, e considerando-se que a conciliação e a mediação são instrumentos efetivos de pacificação social, solução e prevenção de litígios, levantam-se as seguintes hipóteses de pesquisa:

- Há diferenças nas chances de sucesso de conciliação ou mediação com relação ao prazo de duração dos processos.
- Há diferenças nas chances de sucesso de conciliação ou mediação com relação aos valores inicialmente solicitados e negociados.
- Há diferenças nas chances de sucesso de conciliação ou mediação com relação à matéria envolvida.
- Há diferenças nas chances de sucesso de conciliação ou mediação com relação à característica socioeconômica da região;
- Há diferenças nas chances de sucesso de conciliação ou mediação com relação às características das partes.

3. Metodologia

3.1 Descrição e abrangência do universo de pesquisa

O universo da pesquisa consistirá em 100% dos dados disponíveis no contexto digital dos últimos 5 anos da Justiça Estadual dos estados do Ceará, São Paulo, Paraná, Rio de Janeiro e Piauí, disponíveis por meio das decisões publicadas nos respectivos diários oficiais.

3.2 Detalhamento da metodologia de coleta e geração de dados

A mineração de dados é uma tecnologia poderosa que pode ser definida como o processo automatizado de extração de conhecimento e informações úteis, incluindo padrões, associações, ou seja, é uma tecnologia que combina métodos tradicionais de análise de dados com algoritmos sofisticados para o processamento de grandes volumes de dados (VENKATADRI; REDDY, 2011).

Han; Kamber; Pei (2012) descrevem o processo de descoberta de conhecimento como uma sequência iterativa das seguintes etapas:

1. Limpeza de dados: para remover o ruído e dados inconsistentes;
2. Integração de dados: etapa em que várias fontes de dados podem ser combinadas;
3. Seleção de dados: os dados relevantes para a tarefa de análise são recuperados do banco de dados;
4. Transformação: os dados são transformados e consolidados em um padrão apropriado para mineração, realizando operações de resumo ou agregação;
5. *Data mining*: processo essencial em que métodos são aplicados para extrair padrões de dados;
6. Avaliação de padrões: para identificar os padrões verdadeiramente interessantes que representam o conhecimento com base em medidas de interesse, e;
7. Apresentação: etapa em que a visualização e as técnicas de representação do conhecimento são usadas para apresentar conhecimento.

No que se refere ao contexto do projeto, o processo de descoberta do conhecimento anteriormente descrito será aplicado conforme esquema da Figura 1.

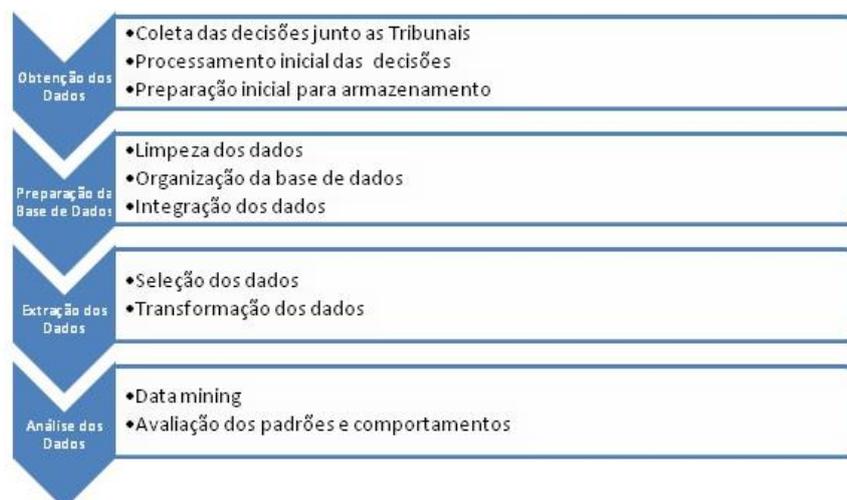


Figura 1 – Etapas para a realização da pesquisa

A etapa de obtenção dos dados acontece antes da realização da limpeza dos dados (primeira etapa do processo de descoberta do conhecimento). Para que haja sucesso em qualquer pesquisa empírica é necessário o acesso a dados confiáveis, em quantidade suficiente para que se minimizem os efeitos de vieses cuja ocorrência pode inviabilizar o resultado final desejado. Nesse aspecto, o presente projeto visa trabalhar com fontes de dados oriundas de informações oficiais públicas. Para tanto é necessário estabelecer duas hipóteses, quais sejam, a possibilidade de acesso dos dados de modo centralizado ou difuso, mediante a colaboração do próprio CNJ ou captura dos mesmos dados de modo programático via sistemas informatizados.

Escritório de Apoio à Pesquisa

Caso haja a disponibilização dos dados de modo centralizado pelo CNJ, ou mesmo de modo difuso diretamente pelos Tribunais alvo da análise, a fase de obtenção de dados poderá ser abreviada ou mesmo suprimida, restando, nesse caso, apenas a organização desses dados de modo a facilitar a consulta e aplicação dos demais métodos adotados. Outrossim, caso haja a necessidade de se obter tais dados diretamente de suas fontes, o trabalho de coleta de dados deverá tornar-se mais complexo. Todavia tal complexidade, hodiernamente, é facilmente resolvida através da tecnologia.

Tendo em vista que atualmente todos os tribunais, estaduais e federais, dispõem de sistemas informatizados de consulta aos andamentos processuais, bem como possuem um sistema de publicações oficiais por meio de diários digitais, é possível acessar uma grande base de dados que servirá de matéria prima da qual a pesquisa se alimentará.

Após a recuperação dos dados, é necessária a realização de um processamento inicial que consiste em converter tais dados em texto puro (extensão de arquivo .txt), na separação das páginas e a organização e padronização dos dados, com a finalidade de preparação inicial para o armazenamento.

Armazenados os arquivos, pretende-se organizá-los em algum sistema de gerenciamento de banco de dados (SGBD) noSQL. Em contraposição aos SGBD relacionais, os noSQL são projetados para armazenamento e processamento de dados em grande escala, principalmente para lidar com enormes quantidades de dados as quais soluções SGBDs convencionais não conseguiram lidar. Eles podem suportar várias atividades, incluindo análises exploratórias e preditivas, transformação de dados em no modelo ETL, gerenciando transações de longa duração ou interorganizações, por exemplo (MONIRUZZAMAN; HOSSAIN, 2013).

Depois de organizados tais dados é possível identificar os processos que se enquadrem no tema proposto filtrando todas as publicações que a eles se relacionem (extração dos dados). A partir das publicações filtradas é possível analisar quais dessas tratam-se de sentenças judiciais de primeira instância, partindo então para a análise do conteúdo desse material, identificando padrões que se alinhem com as hipóteses levantadas. É possível ainda a análise complementar do trâmite de tais processos nas varas originais por meio da captura eletrônica de seus andamentos, o que pode complementar conclusões tais como tempo efetivo de tramitação, assim como o efeito dos atos não publicáveis no deslinde de cada causa, o que possibilitaria, hipoteticamente uma análise do trabalho dos órgãos judiciais envolvidos (análise dos dados).

3.3 Adequação da proposta de pesquisa aos métodos e técnicas de coleta/geração de dados

Considerando o grande volume de dados previsto, bem como a não existência de uma padronização efetiva, a utilização de um SGBD noSQL se apresenta como a mais indicada pela facilidade de extração das informações. A mineração de dados complementa a indicação, uma vez que visa a exploração de grandes volumes de dados, com o objetivo de extrair conhecimentos por meio do reconhecimento de padrões e relacionamento entre variáveis. Além disso, serão utilizados de softwares específicos para análise de dados extraídos.

Considerando assim que esses conhecimentos possam ser obtidos por técnicas comprovadamente confiáveis e validados pela sua expressividade estatística, serão respondidas as perguntas desta pesquisa.

3.4 Técnicas de análise dos dados

A etapa inicial em termos de mineração de dados é a análise descritiva, que busca tanto descrever fatos relevantes, não triviais e desconhecidos dos usuários, como analisar a base de dados, principalmente pelo seu aspecto de qualidade, para validar todo o processo da mineração e seus resultados.

Nesse âmbito, a aplicação de estatística descritiva (proporções, médias, desvios, dentre outros) e gráficos (diagramas de dispersão, diagrama de caixa, histogramas) deverão ser suficientes para obtenção do panorama da fase 1 do projeto.

Em continuidade, poderão ser aplicadas outras técnicas de mineração de dados, tais como Análise de Regras de Associação, Análise de Padrões Sequenciais e Classificação e Predição, Análise de Redes, dentre outras, dependendo dos resultados obtidos previamente.

Escritório de Apoio à Pesquisa

Para tanto, o uso de ferramentas computacionais (MS-Excel) e softwares estatísticos (como o R e o SPSS – *Statistical Package for the Social Sciences*), e outros mais específicos (Power BI, Tableau, Atlas TI, QDA Miner), é imprescindível.

Os dados empíricos coletados na segunda etapa serão colacionados em um relatório descritivo. A partir de então, serão estudados em conjunto com os dados inicialmente obtidos de modo a fornecerem os resultados práticos esperados.

4. Atividades

Para a realização do projeto, pretende-se executar as seguintes atividades, obedecendo o cronograma no Quadro 1. Também será observada a obrigatoriedade dos relatórios parcial e final.

1. Configuração do Servidor

Serão utilizados inicialmente recursos do Internuvem USP. A tarefa inicial é configurar o servidor no tocante a questões de segurança. Também será necessário a instalação do SGBD MongoDB, bem como as ferramentas acessórias.

2. Organização da base de dados

Com o início da coleta dos dados, é necessário realizar a modelagem e estruturação da base de armazenamento. Como mencionado, pretende-se organizá-los em algum sistema de gerenciamento de banco de dados (SGBD) *noSQL* que são projetados para armazenamento e processamento de dados em grande escala. O sistema de gerenciamento escolhido foi o MongoDB.

3. Integração dos dados

Como explicado, os dados provem das publicações dos Tribunais de Justiça dos Estados que compõem o projeto. Os dados não estão padronizados e organizados para armazenamento. Após a realização da etapa 1, buscar-se-á a elaboração de um modelo de dados para promover a integração dos dados e assim facilitar a extração dos dados.

4. Seleção dos dados

Uma vez estruturada, a base de dados estará pronta para a realização de consultas para obtenção dos resultados. Nessa fase serão elaboradas as *queries* de acordo com as necessidades da equipe de pesquisa.

5. Monitoramento, melhorias e suporte para obtenção dos dados

Esse item consiste no contínuo monitoramento de melhorias, contribuindo para o desenvolvimento de consultas padronizadas por meio de uma interface amigável, facilitando a busca de informações.

Quadro 1 - Cronograma previsto

Ativ	Mês 1	Mês 2	Mês 3	Mês 4	Mês 5	Mês 6	Mês 7	Mês 8	Mês 9	Mês 10	Mês 11	Mês 12
1												
2												
3												
4												
5												

5. Resultados Esperados

O principal resultado esperado é fornecer o suporte para a realização da pesquisa apoiada pelo CNJ. A infraestrutura lógica de armazenamento implementada será de fundamental importância para alcançar os objetivos da pesquisa (seção 2.1).

Além disso, a base de dados ficará disponível para outros pesquisadores que queiram extrair dados para realização de suas pesquisas. Adicionalmente, o trabalho pode contribuir para o acesso à informação.

Referências

CNJ. **CNJ Pesquisa**. Disponível em: <<http://www.cnj.jus.br/corregedoriacnj/190-gestao-planejamento-e-pesquisa/18884-cnj-pesquisa>>. Acesso em: 15 maio. 2018.

EDITAL 02/2017. **Edital de convocação pública e de seleção convocação**. Disponível em: <<http://www.cnj.jus.br/files/conteudo/arquivo/2017/09/780a1047c628a90da5a29928f0d4dc01.pdf>>.

Acesso em: 15 maio. 2018.

HAN, J.; KAMBER, M.; PEI, J. **Data mining concepts and techniques, third edition**. 3rd. ed. Waltham, Mass.: Morgan Kaufmann Publishers, 2012.

LAUDON, K.; LAUDON, J. **Sistemas de informações gerenciais**. 11a. ed. [s.l.] Pearson/Prentice Hall, 2015.

LETTIERI, N. et al. Ex Machina: Analytical platforms, Law and the Challenges of Computational Legal Science. **Future Internet**, v. 10, n. 5, 26 abr. 2018.

MONIRUZZAMAN, A. B. M.; HOSSAIN, S. A. Nosql database: New era of databases for big data analytics-classification, characteristics and comparison. **arXiv preprint arXiv:1307.0191**, v. 6, n. 4, p. 1–14, 2013.

VENKATADRI, M.; REDDY, D. L. C. Article: A Review on Data mining from Past to the Future. **International Journal of Computer Applications**, v. 15, n. 7, p. 19–22, 2011.