

## PROJETO DE DISSERTAÇÃO

- Orientadores: Leandro Alvim e Rodrigo Tavares
- Título: Processamento de Linguagem Natural aplicado ao Discurso Jurídico
- Objetivo: Aplicar ferramentas computacionais de processamento de linguagem natural para investigação de textos jurídicos.
- Temas envolvidos: Argumentação Jurídica, Retórica Moral, Plágio, Citação de Precedentes, Processamento de Linguagem Natural
- Introdução ao tema: O fornecimento de razões escritas para fundamentar decisões judiciais, deliberações legislativas e atos administrativos é atividade inerente à legitimação do exercício do Poder num Estado de Democrático de Direito. Portanto, textos são fontes primordiais para compreensão de todo e qualquer procedimento de criação, interpretação e aplicação do Direito.

Por outro lado, os recentes avanços na área de Inteligência Artificial, tais como o desenvolvimento de novas técnicas de processamento de linguagem natural, através do aprendizado de máquina, permitem tratar dados textuais em quantidades antes inacessíveis para leitores humanos. Dessa forma, o uso de novas ferramentas computacionais oferece meios alternativos para a pesquisa de documentos jurídicos, possivelmente, elucidando padrões não distinguíveis em pesquisas textuais de menor escala e, assim, criando conhecimento novo sobre o funcionamento do ordenamento jurídico.

- Descrição sucinta da metodologia:

Nosso objetivo específico consiste de encontrar os valores semânticos que fundamentam uma argumentação jurídica, bem como tendências e contradições dos mesmos. Para tal, elencamos a seguir, algumas etapas da nossa metodologia.

- Selecionar documentos a partir de fontes do governo
- Extrair textos desses documentos
- Aplicar métodos de processamento de linguagem natural para determinar os valores semânticos do conteúdo de cada documento. Algumas métricas como: frequência de palavras; análise de sentimentos; riqueza de vocabulário do discurso; extração de padrões gramaticais e semânticos; e agente e paciente podem ser exploradas.
- Mapear e agrupar conjuntos de dados quanto a sua similaridade em valores semânticos
- Gerar estatísticas acerca dos valores semânticos
- Visualizar os valores semânticos
- Buscar tendências e contradições