## IndIR - Indexing Improving Information Retrieval in the Juris Dataset (Jurisprudence Statements and Thesaurus of the Brazil Federal Court of Accounts - TCU)

Leonardo Pacheco e Marcus Borela https://github.com/marcusborela/ind-ir Projeto Final da disciplina IA368DD 2023S1: Deep Learning Aplicado a Sistemas de Buscas Semana 1 Semana 2 Semana 3 Semana 4 Semana 5 Semana 6 18 19 20 21 22 23 24 25 26 27 28 29 30 31 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 ATRIBUÍDO TAREFA TÉRMINO PROGRESSO INÍCIO Planejamento Plano básico LM 100% 18/5/23 2/6/23 1a Apresentação LM 100% 22/5/23 24/5/23 100% Definir base de dados LM 23/5/23 31/5/23 Definir métricas LM 75% 25/5/23 18/6/23 Construção do Dataset de Indexação (Index-VCE-Juris-TCU) 80% 27/5/23 2/6/23 Carga dos dados do sistema Juris (incluindo indexações) LM 50% Explorar dados do sistema Juris LM 1/6/23 5/6/23 20% Explorar dados de indexações LM 5/6/23 7/6/23 50% 27/5/23 7/6/23 Tratar dados LM Gerar dataset (com grels) 0% 7/6/23 9/6/23 M Validar formato dataset М 0% 7/6/23 9/6/23 Construção do Dataset de Avaliação de Busca (Search-Juris-TCU) Explorar dados de log da busca 20% 14/6/23 1/6/23 0% 8/6/23 Criar queries 5/6/23 Definir prompt LLM para grel LM 0% 8/6/23 12/6/23 Montar base de avaliação (grel) L 0% 8/6/23 18/6/23 Construção do pipeline de indexação 5/6/23 Criar estrutura busca (elastic search, etc) M 0% 4/6/23 Criar estrutura de registro e avaliação (métrica) М 0% 4/6/23 7/6/23 Realizar experimentos: indexador como pipeline de busca LM 0% 7/6/23 21/6/23 Treinar modelo M 0% 12/6/23 21/6/23 Frente pipeline de busca Criar estrutura busca (elastic search, etc) М 0% 14/6/23 15/6/23 Criar estrutura de registro e avaliação (métrica) 0% 14/6/23 16/6/23 М 21/6/23 Realizar experimentos: indexador como ranker LM 0% 14/6/23 Frente Avaliação e Fechamento Relatório LM 16/6/23 28/6/23 0% Preparar apresentação final 0% 28/6/23 LM 24/6/23 0% Publicar datasets com grels (indexação e busca da jurisprudência) LM 4/6/23 28/6/23