

## Aplicação web estilo chatbot para responder perguntas procurando informação em documentos.

Kevin Castro Weitgenant<sup>1</sup>; Tiago Thomsen Primo<sup>2</sup>

<sup>1</sup> Universidade Federal de Pelotas – kcweitgenant@inf.ufpel.edu.br

<sup>2</sup> Universidade Federal de Pelotas – tiago.primo@ufpel.edu.br

### 1. INTRODUÇÃO

Atualmente, a produção de informação é constante, em elevada quantidade e em proporções nunca vistas antes. Estima-se que necessitaremos analisar cerca de 150 trilhões de gigabytes de dados não estruturados em 2025. Dessa forma, existe a necessidade de fazer sentido de toda essa informação em tempo hábil, *Document Understanding* voltados a *Question Answering* se apresentam como uma valiosa ferramenta para esse objetivo.

*Document Understanding* é a habilidade de sistemas de inteligência artificial capazes de processar documentos automaticamente, sendo algumas das possíveis tarefas:

- Document Layout Analysis (DLA) — Processo de identificar e categorizar regiões de interesse em um documento
- Tabular Data Comprehension (TDC) — Compreensão de dados tabulares
- Document Question Answering (DQA) — Responder perguntas de humanos sobre um dado documento

O presente artigo apresenta o desenvolvimento de uma aplicação web em estilo chatbot, onde é possível que os usuários façam perguntas sobre documentos extensos e a aplicação procura passagens relevantes no mesmo, tendo como objetivo facilitar o acesso a essas informações. Para tal, foram utilizados documentos da UFPEL (guia de integralização extensão) e o serviço de Smart Understanding da IBM Cloud, a aplicação web desenvolvida faz comunicação via API com o serviço da IBM e retorna os resultados para os usuários.

### 2. METODOLOGIA

O trabalho consistiu primeiramente na pesquisa de tecnologias nesse sentido, com a tecnologia de *Document Question Answering*.

Após levantada as opções, a solução da IBM foi escolhida. Seguiu-se o processo de aprendizagem de vários conceitos sobre desenvolvimento web.

Em um primeiro momento, foi feita uma análise de projetos semelhantes e então com os conhecimentos adquiridos foi possível o desenvolvimento da aplicação.

### 3. RESULTADOS E DISCUSSÃO

Exemplificando a aplicação em funcionamento, da seguinte forma:

- Acessar a aplicação <https://chatbotcompassagens.herokuapp.com/>
- Uma pergunta sobre o documento foi realizada, nesse caso: “Cursos que podem ser ministrados nas faculdades” (Imagen 1)
- Após apertar ENTER, uma lista de passagens relevantes são retornadas (Imagen 2)

Imagen 1

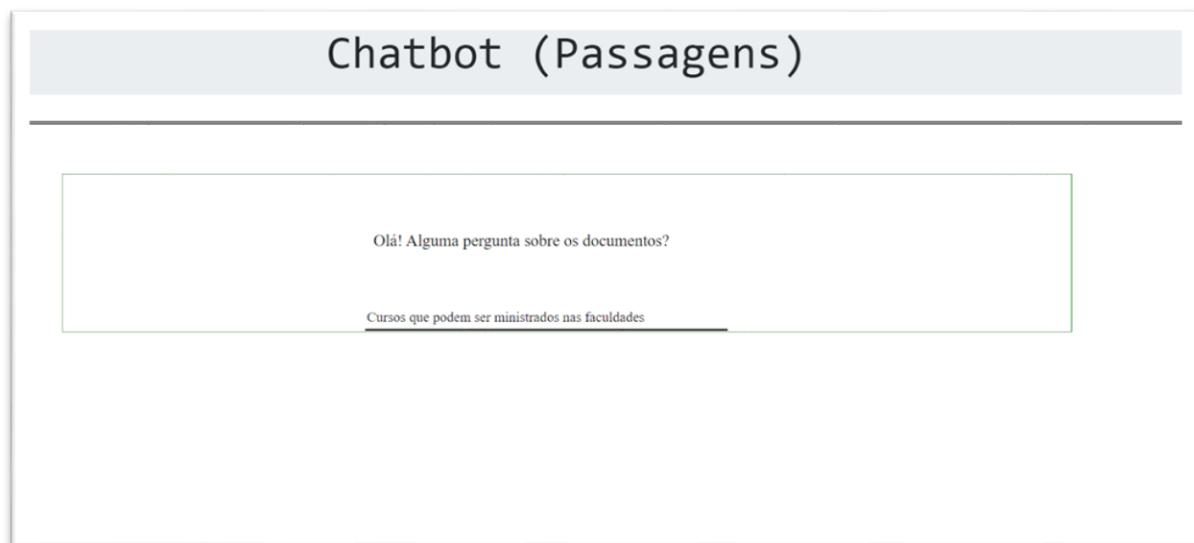


Imagen 2

### Cursos que podem ser ministrados nas faculdades

Art. 69. Nos estabelecimentos de ensino superior podem ser ministrados os seguintes cursos: a) de graduação, abertos à matrícula de candidatos que hajam concluído o ciclo colegial ou equivalente

, da pesquisa e da extensão) e 9 (Intensificar as relações entre UFPel e sociedade), além de outros que o curso possa atender. Assim, já podem ser contempladas as “potencialidades do curso para o atendimento de necessidades e demandas da região”, como indicado no item.

As indicações que seguem contemplam recomendações do que deve constar nestes campos e os conteúdos que informam sobre como se dá no curso a formação em extensão. Portanto, as informações deste capítulo

constituem orientações gerais que podem ser seguidas pelos

É possível constatar que a aplicação retorna em sua primeira passagem(das três retornadas) informações relevantes a pergunta em questão.

## 4. CONCLUSÕES

A aplicação se mostrou satisfatória para os objetivos almejados. Algumas limitações foram observadas(extração de informação de tabelas por exemplo) devido a utilização da versão gratuita do Document Understanding do Watson Discovery.

Devido a ser o primeiro contato do autor na época com desenvolvimento web, a interface da aplicação não foi o foco principal e certamente deve ser aprimorada caso ocorra eventualmente uma entrega para o público.

## 5. REFERÊNCIAS BIBLIOGRÁFICAS

Medium. 2022. *Awesome Document Understanding / AI Document Processing*. [online] Available at: <<https://medium.com/responsibleml/awesome-document-understanding-ai-document-processing-c609ffed2e15>> [Accessed 22 August 2022].