

Proposta de assistente virtual inteligente para estudantes das instituições federais de ensino profissional

Paulo A. Sabaini¹, Bruno G. R. Clemente², Bruno C. Coutinho³

¹Café Meridiano Indústria e Comércio Ltda - Colatina - ES - Brazil

²Matemática, Física e Informática do Instituto Federal do Pará - Paragominas/PA - Brazil

³Instituto Federal do Espírito Santo campus Serra (IFES) - Serra - ES - Brazil

pasabaini@gmail.com, bruno.clemente@ifpa.edu.br, bccout@ifes.edu.br

Abstract. *A federal education institute has a variety of information concerning its operation, distributed on several pages of its web portal and documents of the type pdf. Often accessing this information can become a frustrating task, even using the available search engines. The purpose of this work is to develop a prototype of a chatbot, that allows raising the possibilities and challenges in the construction of an intelligent system, that solves primarily the doubts regarding the pedagogical management of an educational institute, with the knowledge base formed through textual data already available on the IF portal. The first results were promising.*

Resumo. *Um instituto federal de ensino possui uma variedade de informações, concernentes ao seu funcionamento, distribuídas em várias páginas do seu portal web e documentos do tipo pdf. Muitas vezes o acesso a essas informações pode se tornar uma tarefa frustrante, mesmo utilizando os mecanismos disponíveis de busca. A proposta deste trabalho é desenvolver um protótipo de chatbot, que permita levantar as possibilidades e desafios na construção de um sistema inteligente, que resolva primariamente as dúvidas em relação à gestão pedagógica de um instituto de educação, com a base de conhecimento formada através de dados textuais já disponíveis no portal do IF. Os primeiros resultados se mostraram promissores.*

1. Introdução

A mineração de dados textuais é um método interdisciplinar que envolve as áreas de recuperação de informação, aprendizagem de máquina, estatística, linguística e a mineração de dados propriamente dita [Santos et al. 2014]. O Processamento de Linguagem Natural (PLN) e a mineração de dados caminham juntas na obtenção de informações (navegação, organização e descoberta inteligente) [Ferrari 2018] em bases de dados não estruturados, como é o caso dos textos escritos em linguagem natural, uma fonte abundante na internet.

Os institutos federais de educação profissional (IF) possuem um conjunto de regras e padrões que são diariamente publicados e atualizados em seus portais na Internet, através de páginas distribuídas a partir de muitas seções e links, ou mesmo via documentos no formato “pdf”. Essas informações montam uma grande base de conhecimento que, muitas vezes não possui um acesso fácil e uma linguagem acessível

2. Revisão de Literatura

O PLN tem como meta produzir ferramentas que compreendam a língua, ou seja, habilitar um dispositivo eletrônico a processar a mesma linguagem que os humanos usam no dia a dia [Comarella and Café 2008]. A avaliação de frases pode ser realizada com o apoio de recursos de PLN e regras linguísticas que permitem a identificação dos elementos importantes das respostas, que serão consultadas em uma ontologia de domínio para identificar a sua correção [Pereira and Rigo 2013].

Ainda em se tratando de PLN, a mineração de textos pode ser empregada na obtenção de informações importantes para o contexto em que o sistema inteligente está sendo aplicado. Considerando a gama de textos disponíveis / publicados na internet e, no contexto deste projeto, mais especificamente no portal do IFES, traz um problema de dispersão da informação entre várias fontes, gerando uma sobrecarga, que faz com que o processo de busca de uma informação relevante seja mais complexo, geralmente envolvendo a análise de uma grande quantidade de dados textuais. A mineração de textos oferece um conjunto de métodos que permite a navegação, organização e descoberta inteligente de informação em bases de dados não-estruturados [Santos et al. 2014].

Sistemas de Pergunta-Resposta (PR) são programas que recebem como entrada de dados uma pergunta em linguagem natural e, mediante pesquisa em alguma base de dados, retornam a resposta esperada. Os primeiros sistemas interativos de perguntas e respostas em português foram o Elektra [Leonhardt et al. 2003] e o “Pergunte!” [Rabelo and Barros 2004].

Observa-se o crescimento nesta última década das tentativas de pesquisadores, desenvolvedores e empresas importantes do mercado da tecnologia da informação e comunicação (TIC), em se otimizar problemas do cotidiano através das capacidades linguísticas associadas ao computador: escrita e produção de textos, sistemas de busca em dados não estruturados, tradução, indexação de documentos e sistemas interativos [Comarella and Café 2008]. Tais sistemas interativos de PR surgem, muitas vezes, do desenvolvimento de *chatterbots* (ou simplesmente *chatbot*, ou apenas *bot*).

Alguns exemplos do uso de chatbots em instituições brasileiras de ensino podem ser citados: a Prof^a Elektra para o ensino de física e redes de computadores [Leonhardt et al. 2003], o Esteban como auxiliar no ensino de língua estrangeira [Ferreira and Uchôa 2006], o Tical usado na divulgação do atlas linguístico do Brasil [Moreno et al. 2015], o Doctrina enquanto um tutor inteligente do Instituto Federal do Triângulo Mineiro (IFTM), construído a partir de técnicas bio-inspiradas [Lima et al. 2017], o CVChatbot que é integrado ao *Moodle* e capaz de interagir com os usuários, enviando notificações sobre módulos criados pelos professores e mensagens dos fóruns [Dehon et al. 2018] etc.

3. Proposta

O *chatbot* aqui descrito tem como objetivo auxiliar o acesso dos alunos às informações específicas de sua instituição de ensino, ajudando-o à filtrar aquelas existentes em uma grande quantidade de documentos disponíveis nas instituições, com o uso de ferramentas que já se fazem presentes no dia a dia dos alunos, em linguagem

descontraída e acessível. Esses documentos muitas vezes não estão organizados, não são apropriadamente formatados e se encontram em uma linguagem de difícil acesso e filtragem.

Não apenas o *chatbot* auxilia o aluno a encontrar as informações que deseja, ele também auxilia os profissionais da educação, de funcionários dos setores específicos, até docentes, que muitas vezes ficam sobrecarregados tendo que atender pessoalmente cada aluno que tem uma demanda específica de uma informação que, em teoria, já se encontra disponível.

O *chatbot* também deve ser capaz de receber novos documentos em PDF para atualizar sua base de conhecimento, permitindo que o *bot* aprenda novos assuntos, sendo capaz de extrair as possíveis perguntas e as respostas que atenda à expectativa do aluno.

O protótipo apresentado neste trabalho representa um recorte da proposta completa: ele é capaz de responder perguntas sobre o ROD dos cursos técnicos do IFES. Sua base de conhecimento foi construída a partir do levantamento das questões mais frequentes dos alunos desses cursos (ensino médio), aquelas feitas rotineiramente aos profissionais de educação que são responsáveis por atender esses alunos e tirar suas dúvidas sobre o conteúdo do ROD.

A Figura 2 descreve a estrutura interna da versão 1.0 do *bot ifes.talk* (Sabaini, 2019), fruto de uma iniciação científica e de um trabalho de conclusão de curso, ambos realizados por um mesmo aluno da graduação em Sistemas de Informação do IFES campus Colatina. O *ifes.talk* também já foi registrado no INPI³. O processo de assistência virtual se inicia com a interação do usuário através de uma interface de diálogo (*chat*), o canal de entrada de dados por onde o aluno pode elaborar sua pergunta ao assistente. Essa requisição é recebida pelo módulo de PLN, desenvolvido na plataforma *Dialogflow*⁴, que usa o conceito de intenções para identificar o que o usuário realmente deseja com a frase digitada. Tal intenção é transmitida ao *back-end*, responsável por selecionar uma resposta adequada e devolvê-la ao usuário, repetindo o processo quantas vezes for necessário.

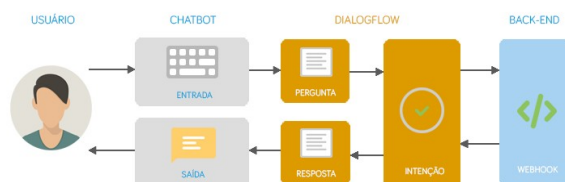


Figura 2. Arquitetura interna atual do *ifes.talk*

O *chatbot* possui os seguintes componentes:

- **Usuário:** alunos e servidores do IFES que farão o diálogo com o *chatbot*;
- **Chat:** são os aplicativos de mensagens que atuam como *front-end* do *bot*;
- **Dialogflow:** ferramenta *cloud-based* de PLN utilizada no projeto;
- **API:** interface com o *back-end* para consulta de respostas e também

³ <http://www.inpi.gov.br/>

⁴ <https://dialogflow.com/>

responsável pelo funcionamento do painel administrativo;

- **Banco de dados:** armazena as respostas, usuários, avaliações e informações para a geração de estatísticas de uso;
- **Painel administrativo:** *website* onde os administradores podem gerenciar as respostas do *chatbot* e visualizar as avaliações e estatísticas de uso, facilitando a utilização e atualização do *bot* por usuários leigos.

A interação com o *bot* se dá por meio de um aplicativo de mensagens, atualmente com duas opções: o *Facebook Messenger* e o *Telegram*, ferramentas que já se fazem presentes nos *smartphones* de muitos alunos das áreas de tecnologia. Ao enviar uma mensagem para o *chatbot*, esta é redirecionada automaticamente para o mecanismo de PLN utilizado, que se encarrega de encontrar a resposta que melhor responda a pergunta feita, que então é enviada de volta para o aplicativo de mensagens.

O painel administrativo da aplicação é acessado pelo administrador do sistema e pelos usuários cadastrados. Esse painel acessa a mesma API, assim, compartilhando os dados em tempo real com o *chatbot*.

Para o desenvolvimento da API e do painel administrativo da aplicação foi utilizada a linguagem de programação Python com o *micro framework web* Flask⁵, lançado em 2010 e que é amplamente utilizado em vários projetos ao redor do mundo.

Na criação da interface *web* do painel administrativo foi utilizado HTML e CSS junto com o *framework Bootstrap*, criando uma interface responsiva. Foi ainda utilizado *Javascript* para a manipulação de informações nas páginas *web* com as bibliotecas *jQuery* e *Chart.js*, sendo esta última utilizada para a criação dos gráficos.

Para o armazenamento dos dados da aplicação foram utilizados dois sistemas diferentes de banco de dados: No ambiente de testes e desenvolvimento foi utilizado o *SQLite*, uma biblioteca de código aberto (*open source*) desenvolvido na linguagem C que permite a disponibilização de um pequeno banco de dados na própria aplicação, sem a necessidade de acesso a um SGBD separado. No ambiente de produção foi utilizado o *PostgreSQL*, normalmente chamado de *Postgres*, um SGBD do tipo objeto-relacional (ORDBMS) com ênfase em extensibilidade e em padrões de conformidade. Como um servidor de banco de dados, sua principal função é armazenar dados de forma segura, apoiando as melhores práticas, permitindo a recuperação dos dados a pedido de outras aplicações de *software*.

A proposta deste trabalho de pesquisa é construir um assistente virtual autônomo escalável e flexível, que consiga atualizar sua base de informações através de documentos, páginas disponibilizadas no portal do IFES e seus subdomínios.

4. Resultados e discussões

Na Figura 3 tem-se um exemplo de conversa com o *chatbot*. Na tela da esquerda o *bot* se apresenta e pergunta se o usuário deseja iniciar a conversa, que após iniciada permite o mesmo exibir as opções disponíveis para auxiliar o usuário em encontrar o tema desejado, bastando selecionar uma das opções para dar sequência à conversa. Contudo é possível também realizar uma pergunta diretamente, sem o auxílio das respostas rápidas (*quick replies*).

⁵ <https://palletsprojects.com/p/flask/>

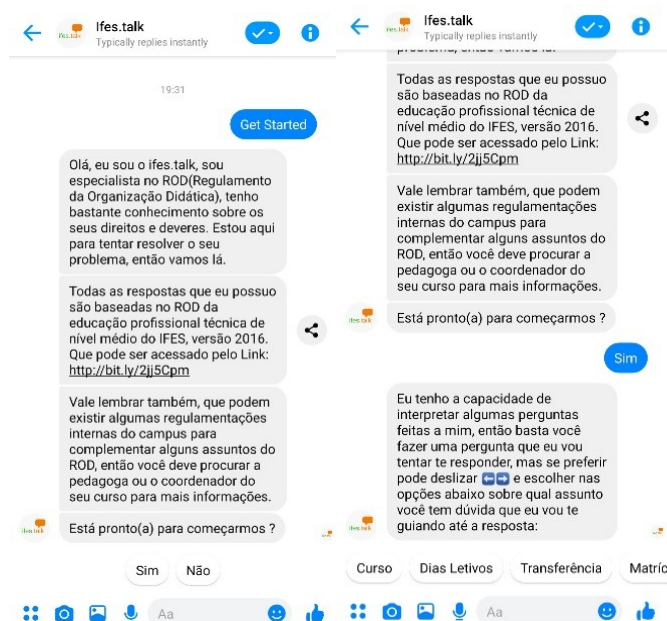


Figura 3. Início da conversa com o *ifes.talk*

O painel administrativo foi construído com o objetivo de facilitar o gerenciamento do *bot* por usuários leigos. Uma de suas funcionalidades é a visualização e alteração das mensagens relacionadas ao ROD. Na Figura 4 é possível visualizar a página onde as mensagens são listadas, sendo possível realizar a filtragem de mensagens, a partir do menu superior, pelo conteúdo da resposta e pelo artigo relacionado a ela. Pelo painel ainda é possível visualizar as estatísticas de uso do *chatbot*, os comentários e avaliações feitas pelos usuários sobre o uso da ferramenta, além de gerenciar aqueles que poderão acessar o painel.

#	Mensagem	Artigo
1	Bom, a transferência para o IFES deve ser feita por meio do Edital de Transferências que é publicado no site do IFES, no Link: https://colatina.ifes.edu.br/editalis &Nesta página, basta buscar pelo edital e nele estarão todas as informações necessárias para a realização da transferência.&Para mais informações, basta consultar o artigo 21 § 2º do ROD.	21 x § 2º
2	Para se transferir para outra escola, o seu representante deverá ir até a CRA (Coordenadoria de registro acadêmico), e pedir a sua transferência.& E então, em até 20 dias úteis os documentos necessários para a transferência ficarão disponíveis, são eles o seu histórico parcial e a ementa das matérias que você já cursou.&Para mais informações, basta consultar o artigo 65 do ROD.	65 x
3	O líder da turma deve participar da reunião pedagógica intermediária.&Para mais informações, basta consultar o artigo 90 do ROD.	90 x
4	Cada turma deve ter um aluno e professor representantes, que são eleitos por meio de uma votação.&Para mais informações, basta consultar o artigo 100 § 2º do ROD.	100 x § 2º
5	A mudança de modalidade (Presencial/Distância) pode ser requerida uma única vez por curso e deve ser feita dentro do prazo descrito no calendário acadêmico.&A mudança só será executada se houver vagas disponíveis.&Os documentos necessários para a realização são: o histórico escolar e a ementa do curso.&Para mais informações, basta consultar o artigo 45 e 46 do ROD.	45 x 46 x

Figura 4. Configuração das mensagens através do painel administrativo

4.1. Avaliação dos alunos

Para a avaliação do protótipo construído neste trabalho, foram reservadas duas aulas de uma turma do primeiro ano do curso técnico integrado de informática para a internet do IFES campus Colatina. No total, 39 alunos participaram da avaliação, sendo divididos em dois grupos, um com 20 alunos e outro com 19 alunos. Primeiramente, o *bot* foi apresentado para os alunos. Em seguida foi disponibilizado um intervalo de 15 minutos

para que os mesmos pudessem utilizar a ferramenta de forma livre, para entender seu funcionamento e se familiarizar com ela. Após este período foi dado início a avaliação que consistia na realização de sete perguntas pré-selecionadas ao *bot*, onde era permitido a efetivação das perguntas com outras palavras para testar a capacidade de compreensão da ferramenta. Antes de realizarem as perguntas foi disponibilizado um link para o ROD dos cursos técnicos e foi instruído que, para cada uma das perguntas, o aluno deveria procurar a resposta para ela primeiramente no ROD e, em seguida, realizar a pergunta para o *bot* para que assim pudesse ser constatada a diferença entre as duas abordagens. No final, foi instruído aos alunos que respondessem 10 perguntas sobre a experiência de uso.

Sobre a utilização do *bot* para esclarecimento de futuras dúvidas, 87,2% dos alunos responderam que usariam a ferramenta em possíveis questionamentos. Ainda sobre a facilidade de uso, a qualidade das respostas e a simplicidade da linguagem do chatbot, 97,4% classificaram como muito bom ou excelente. Essa mesma porcentagem de alunos respondeu que o *bot* seria a primeira opção no caso de futuros questionamentos.

5. Conclusão

O presente trabalho apresentou a proposta de criação de um *chatbot* para apoiar o discente do IFES em relação as suas dúvidas sobre as rotinas acadêmicas, assim como um protótipo cujo objetivo é facilitar a busca por informações no ROD, guiando o aluno até uma resposta esperada, com linguagem simples e clara, além de se tratar de uma ferramenta intuitiva e em linguagem natural, sempre disponível via ferramentas como o *Facebook Messenger* e o *Telegram*.

Também foi desenvolvido um painel administrativo *web* com o objetivo de facilitar a administração do *chatbot*, por membros do projeto e por servidores do IFES, mesmo sem conhecimento de tecnologia, visto que a plataforma é de fácil utilização. O painel administrativo apresenta informações úteis que podem ser utilizadas para a tomada de decisões quanto ao ROD e aos alunos.

Como trabalhos futuros para crescimento da ferramenta, podem-se citar: alimentação automática da base de conhecimento do *bot* através de documentos disponíveis no portal do IFES (ex.: páginas *web*, pdfs, etc...), estender seu conhecimento através de entrevistas com especialistas dos setores administrativos via técnicas de aprendizagem de máquina, ajudar no preenchimento de formulários para dar entrada em processos de requisição geral, integração com outros sistemas do IFES como o Q-Acadêmico que gerencia os diários das disciplinas, pautas e notas dos alunos, e outros. Uma outra ideia futura é permitir que o *ifes.talk* possa ser customizado para atender outros IFs, claramente, em conformidade com suas especificidades.

Referências

- Comarella, R. L. and Café, L. M. A. (2008) “Chatterbot: conceito, características, tipologia e construção”, *Informação & Sociedade: Estudos*, v. 18, n. 2, p. 55–67.
- Dehon, P., Silva, A., Inocêncio, A. C., Castro, C., Costa, H. and Júnior, P. P. (2018)

“CVChatbot: Um Chatbot para o Aplicativo Facebook Messenger Integrado ao AVA Moodle”, Anais do XXIX Simpósio Brasileiro de Informática na Educação (SBIE), 1623-1632, <https://doi.org/10.5753/cbie.sbie.2018.1623>.

Ferrari, A. (2018) “Natural Language Requirements Processing: from Research to Practice Applications”, Gothenburg, Sweden: [s.n.], p. 536–537.

Ferreira, L. P., and Uchôa, J. Q. (2006) “Desenvolvimento de um chatbot para auxiliar o ensino de Espanhol como Língua Estrangeira”, Revista Bazar: Software e Conhecimento Livres, 1, 21–32, <http://repositorio.ufla.br/jspui/handle/1/9629>.

Leonhardt, M. D., Castro, D. D. de, Dutra, R. L. D. S., and Tarouco, L. M. R. (2003) “ELEKTRA: Um Chatterbot para Uso em Ambiente Educacional”, Renote, 1(2), 1–11, <https://doi.org/10.22456/1679-1916.14336>.

Lima, D. A., Oliveira, C. C., Pestili, L. C., Silva, E. C., Bezerra, M. A. S. and Lima, H. A (2017) “Uma proposta de sistema de aprendizagem com conteúdo gamificado e com reforço guiado por algoritmos bio-inspirados”, Anais do Computer on the Beach.

Moreno, F., Manfio, E., Barbosa, C. R., and Brancher, J. D. (2015) “Tical: Chatbot sobre o Atlas Linguístico do Brasil no WhatsApp”, Anais Do XXVI Simpósio Brasileiro de Informática Na Educação (SBIE), 279–288.

Pereira, F. R. and Rigo, S. J. (2013) “Utilização de processamento de linguagem natural e ontologias na análise qualitativa de frases curtas”, Novas Tecnologias na Educação, v. 11, p. 1–10.

Rabelo, J. and Barros, F. (2004) “Pergunte! uma interface em português para pergunta-resposta na web”, Dissertação de Mestrado, Departamento de Computação, Universidade Federal de Pernambuco, Pernambuco, Brasil.

Russel, S. and Norvig, P. (2013) “Inteligência Artificial”, Prentice Hall, 3ª edição.

Sabaini, P. A. (2019) “Proposta de assistente virtual inteligente no apoio ao discente do Ifes em relação às rotinas acadêmicas”, Trabalho de Conclusão de Curso, Colatina - ES.

Santos, R. E. S., Correia-Neto, J. S., Souza, E. P. R., Magalhães, C. V. C. de and Vilar, G. (2014) “Técnicas de processamento de linguagem natural aplicadas ao processo de mineração de textos: resultados preliminares de um mapeamento sistemático”, Revista de Sistemas e Computação, v. 4, p. 116–125.