

## **PROCESSAMENTO DE LINGUAGEM NATURAL (PLN): EVOLUÇÃO, DESAFIOS E PERSPECTIVAS FUTURAS PARA OS CHATBOTS**

José Roberto Madureira Junior<sup>1</sup>, Adani Cusin Sacilotti<sup>2</sup>; Reginaldo Sacilotti<sup>3</sup>

**Resumo** - Este artigo trata da utilização de PLN e *Chatbots* demonstrando sua disseminação impulsionada pela tendência de mercado, bem como pelo avanço tecnológico, que possibilita a solidez, trazendo uma visão histórica e abrangente, de implementação e desafios na área, além das tendências futuras no desenvolvimento de *Chatbots*. Com o conhecimento transmitido ao longo deste artigo foi possível ingressar no contexto do PLN e abordar pontos satisfatórios para a escolha da forma ideal de implementação para um ramo de negócio específico. Desta forma, buscou-se oferecer uma visão geral dos *Chatbots*, que têm sido utilizados para auxiliar os usuários de diferentes áreas a superar diversos desafios.

**Palavras-chave:** Inteligência Artificial, PLN, Chatbots, Agilidade.

**Abstract** - This article deals with the use of NLP and Chatbots, demonstrating their dissemination driven by the market trend, as well as by technological advances, which enable solidity, bringing a historical and comprehensive view of implementation and challenges in the area, in addition to future trends in the development of Chatbots. With the knowledge transmitted throughout this article, it was possible to enter the context of the PLN and address satisfactory points for choosing the ideal form of implementation for a specific line of business. In this way, we sought to offer an overview of Chatbots, which have been used to help users from different areas to overcome various challenges.

**Keywords:** Artificial Intelligence, NLP, Chatbots, Agility.

### **1 Introdução**

O desenvolvimento da humanidade levou a criação de formas de se comunicar, entre elas, a escrita, na qual os seres humanos se tornaram excelentes em ler e interpretar conteúdos de forma rápida e consistente. Compreender o tom, a intensão, a gíria e as abreviaturas é uma tarefa simples para qualquer nativo da língua, processando tanto o texto escrito quanto a fala, no entanto, no que diz respeito as máquinas que processam uma linguagem, o contexto é muito diferente. Na década de 1950, profissionais da área de Tecnologia da Informação (TI) iniciaram o uso de softwares para analisar e processar componentes textuais, partes de discursos, sentimentos e uma grande

---

<sup>1</sup> Pontifícia Universidade Católica de São Paulo - madujr@gmail.com

<sup>2</sup> Faculdade de Tecnologia de Jundiaí – adanics@uol.com.br

<sup>3</sup> Faculdade de Tecnologia de Jundiaí - rsac@terra.com.br

gama de atividades que compõem um texto. Contudo, ainda tem sido um desafio para profissionais de TI (MARKIEWICZ; ZHENG, 2020, p. 11).

Desde o advento do *Watson*, da IBM, a promessa de que máquinas terão a capacidade de interpretar a linguagem natural tem se aproximado da realidade, motivada por recentes avanços nessa área de estudo. Atualmente, vivemos na era das mídias sociais na qual se obtém oportunidade de *insights* de milhões de palavras produzidas todos os dias. Nesse contexto, surgem novas ferramentas que tornam possível que desenvolvedores criem de forma rápida, modelos para entender as palavras informadas na área de interesse.

Um estudo da IBM do ano de 2017 relatou que 90% dos dados do mundo foram criados nos últimos dois anos e que 80% desses dados estão em formato desestruturado (MARKIEWICZ; ZHENG, 2020, p. 11). Dados no formato desestruturado possuem valiosos *insights*, contidos em planilhas de pesquisa, e-mails e até transcrições de reuniões, podendo ser extremamente úteis se forem identificados, interpretados e utilizados. Desta forma, quando os *insights* de dados desestruturados passam a ser utilizados, abrem infinitas possibilidades para melhorar o processo de negócio, reduzindo custos e aprimorando produtos e serviços.

Por outro lado, interpretar esses dados desestruturados é uma tarefa de alta complexidade. Assim, percebe-se que o PLN gerado por seres humanos consiste em um desafio difícil de ser resolvido por um algoritmo projetado para um determinado propósito.

O presente artigo apresenta uma visão geral do PLN, assim como sua ascensão, implementação em ambiente de *Cloud Computing*, interface e tendências que contribuem para empresas que desejam definir uma estratégia de adoção do PLN em aplicações corporativas. Outro ponto ressaltado pelo artigo é apresentação da subárea de *Chatbots*, que são tendências de mercado e apresenta um estudo das perspectivas futuras da PLN.

## 1.1 Histórico

Um dos mais antigos objetivos da interação humano-computador consiste na criação da possibilidade de que pessoas interajam com máquinas através de um diálogo natural, como aconteceria na interação entre dois humanos. Tal situação abriu espaço para a criação e o desenvolvimento de inúmeros recursos e ferramentas pautados nos princípios da Inteligência Artificial (IA), objetivando a melhoria na linguagem, comunicação e interação entre os usuários e as máquinas (JOHRI et al., 2021).

Uma máquina, chamada *Enigma*, foi construída pelos alemães durante a Segunda Guerra Mundial, para ser utilizada na criptografia de mensagens, como estratégia para o exército alemão. Mais tarde, em 1946, na Grã-Bretanha, era apresentada a *Colossus*, uma máquina com capacidade de decifrar os códigos gerados pela *Tunny* (codinome dado à *Enigma* pelos britânicos). Com a capacidade de decifrar os códigos gerados pela *Enigma*, a guerra teve uma grande virada, pois os britânicos puderam conhecer a posição dos militares

alemães, condições do exército, suas estratégias, forças e, assim, puderam tomar ações preventivas.

Ainda no contexto mencionado, durante a Segunda Guerra Mundial, os avanços no campo da criptografia do governo britânico foram estabelecidos em *Bletchley Park*, lugar onde *Alan Turing*, junto com outros agentes de inteligência, decodificavam as mensagens militares alemãs, fato que levou a conquista pioneira durante a Segunda Guerra Mundial e marcou a primeira onda de avanços no campo da computação moderna (JOHRI et al., 2021).

Assim, na década 1950, o Teste de Turing foi proposto por Alan Turing visando analisar se um computador pode pensar como um humano ou não. Este teste não tinha o objetivo de solucionar problemas, mas identificar se a máquina pode pensar da mesma forma que os humanos pensam. Porém, em 1957, não havia como incorporar a gramática nas máquinas e fazê-las compreender o seu verdadeiro sentido. A evolução no campo do PLN aconteceu no mesmo ano, quando Noam Chomsky iniciou as estruturas sintáticas. Chomsky trabalhou no refinamento do conjunto de regras para incorporar a linguística baseada na gramática universal. Posteriormente, Charles Hockett descobriu várias desvantagens como considerar a linguagem como uma estrutura bem definida, estável e em um sistema formal.

Uma das primeiras áreas do PLN foi a tradução automática. O propósito é focado no desenvolvimento de aplicações para realizar a tradução de fala ou texto de um idioma natural para outro. Um exemplo desse tipo de aplicação foi um projeto entre a *IBM Company* e a Universidade de Georgetown, em 1954, que, em experimentos, traduziu com sucesso frases russas para o inglês. No entanto, o processo realizado pela aplicação era mais lento do que o esperado, reduzindo drasticamente o financiamento do projeto (JOHRI et al., 2021).

No final da década de 1960, foi desenvolvida por Terry Winograd, no MIT (*Massachusetts Institute of Technology*), a chamada *SHRDLU*, o primeiro programa de computador capaz de mover objeto, determinar o estado atual e lembrar nomes executando todas essas tarefas no ambiente *blocksworld*. Sendo assim, tornou-se a primeira grande conquista da IA a atrair a atenção dos pesquisadores para otimizar essa tecnologia, mas seu sucesso se limitou quando sua aplicação não obteve êxito com situações do mundo real, ambíguas e complexas.

No ano de 1969, Roger Schank introduziu o importante conceito de *tokens* que proporcionou uma melhor compreensão do significado das frases. Esses *tokens* incluem tempo, local, ações e objetos do mundo real. Com seu uso, as máquinas produzem melhores *insights* sobre as coisas que estão acontecendo na frase e os objetos envolvidos. Neste contexto, todas as regras criadas auxiliavam uma máquina a compreender o significado de uma frase (JOHRI et al., 2021).

Já no início da década de 1970, William Woods introduziu o conceito de *Augmented Transition Networks* (ATNs<sup>4</sup>) para representação de linguagem natural, conceito que surge com um palpite quando a frase fornecida está em uma situação ambígua.

Grande parte das abordagens utilizadas até então para entender as linguagens naturais baseava-se em regras manuscritas onde a máquina costumava realizar combinações. Mas, na década de 1980, o *machine learning* ganhava destaque como outro campo emergente da computação, de modo que os algoritmos forneceram uma forma mais avançada para interpretar a ambiguidade e justificar com evidências as tomadas de decisões. Algoritmos como o de árvores de decisões utilizam regras como *if then* para produzir o resultado ideal, e os probabilísticos respaldam a decisão, pela máquina escolhida, oferecendo confiança no resultado (JOHRI et al., 2021).

Atualmente, uma alteração foi realizada no domínio do *deep learning* para possibilitar a realização de tarefas difíceis de executar usando regras e critérios fixos. Quando se trata de PLN, ainda não há como resolver a ambiguidade da linguagem, visto que uma única palavra pode ter diferentes significados em diferentes contextos, não sendo possível escrever uma regra ou utilizar uma árvore de decisão para representar todos os significados possíveis. Com o *deep learning*, esse problema pode ser resolvido de forma eficiente, posto que o programador não necessita fornecer regras, pois o algoritmo faz a dedução conforme o mapeamento do processo com as entradas e saídas.

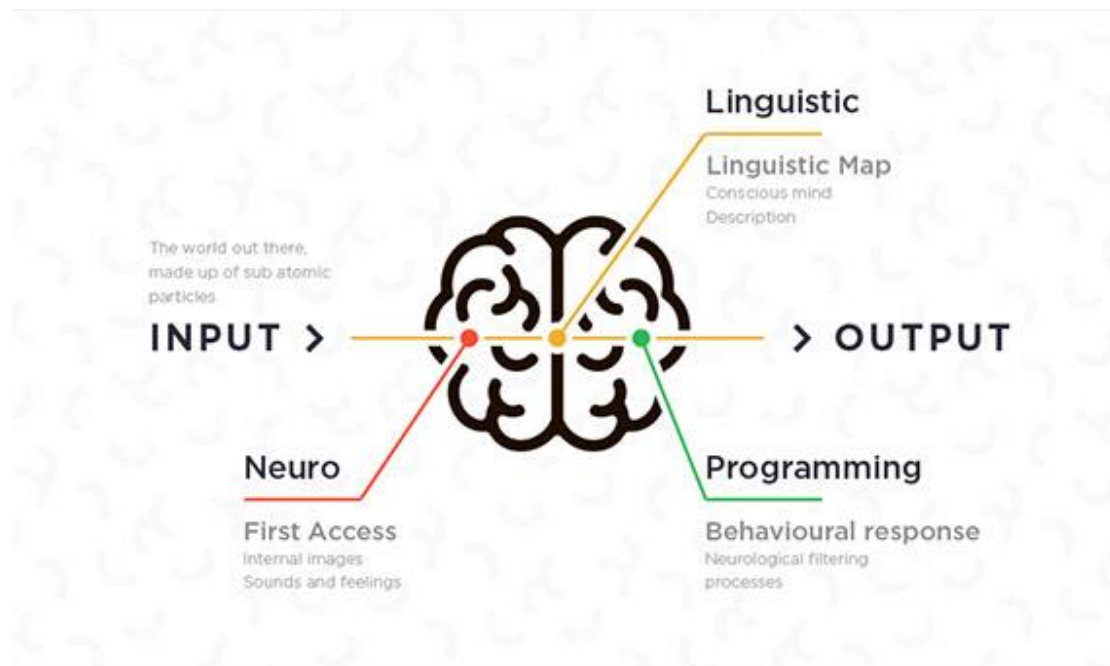
## 1.2 Funcionamento

O processo desenvolvido pela PLN em um computador é similar à aprendizagem de um novo idioma. Dessa forma, primeiro se aprende o vocabulário e, depois, como criar frases simples, sendo que a sintaxe e os conectores exercem um papel fundamental para auxiliar um indivíduo a entender o que está lendo ou ouvindo. À medida que o indivíduo avança no aprendizado de uma nova língua, pode começar a entender composições mais complexas e até gírias ou linguagem técnica (MORGAN, 2018).

---

<sup>4</sup> Teoria que utiliza autômatos de conjuntos finitos com o conceito de recursão para chegar ao estado final e concluir o significado da frase de acordo com as informações disponíveis.

**Figure 1 - Criação de Novo Código PLN.**



**Fonte:** Gdoura (2020).

A Figura 1 apresenta a criação de um código PLN, onde a linguística computacional faz a combinação da IA, *machine learning* e linguística. Cada palavra é dividida para dar entrada no modelo, procurando-se por palavras-chave, como local, horário, canais ou similares. Em seguida, é realizada a verificação da sintaxe, com concentração em verbos, conectores e negações (MORGAN, 2018).

O algoritmo realiza primeiro a análise lexical, classificando a entrada de acordo com a parte do discurso. Posteriormente, é realizado o aprofundamento no nível da análise sintática para realizar a análise das partes da frase. Em alguns aplicativos, outras etapas se fazem necessárias, como análise de sentimento, para possibilitar a compreensão do estado do emissor.

O PLN se baseia em grafos contendo sinônimos, merônimos, antônimos e muito mais. É dessa forma que máquinas podem compreender outro humano e até responder de maneira que não soe robótica. Nesse sentido, o PLN obtém um conjunto de dados de treinamento como entrada a partir do qual deve ser realizado o aprendizado, onde existem três métodos de treinamento diferentes (MORGAN, 2018):

- Aprendizagem guiada baseada em regras, sendo muito semelhante à programação de computador clássica (loops/ if-then-else/ while);
- Aprendizagem estatística construída sobre a resposta mais provável;
- Aprendizado não supervisionado, adequada para sistemas que já estão treinados e pretendem obter melhorias na interação.

### 1.3 Implementação

Existem várias estratégias para implementar a PLN em aplicações corporativas. As primeiras formas de implementar merecem ser destacadas em projetos de software de código aberto, e pode-se destacar o *Apache NLP*, o *Stanford CoreNLP*, o *NLTK* para *Python* e o *SintaxeNet* (MARKIEWICZ; ZHENG, 2020, p.18).

Embora as soluções anteriormente dispostas sejam as mais corriqueiras na construção de aplicações PLN, há uma coleção de bibliotecas de código aberto que podem ser utilizadas nas mais diversas linguagens. Por exemplo, para o desenvolvedor que utiliza a *Ruby*, pode-se encontrar uma coleção de pequenas bibliotecas *Ruby NLP*, e o mesmo pode ser aplicado ao *PHP*, dado que existe a *PHP NLP Tools*, onde não é necessário reinventar algoritmos, basta apenas utilizá-los.

Uma alternativa ao uso de bibliotecas para PLN reside na utilização de APIs, que olhando sob a perspectiva custo-benefício, pode fazer sentido para aplicativos corporativos que recorram a serviços de terceiros. Recentemente, várias empresas fornecem APIs no formato de serviços, entre elas está a IBM, com o *IBM Watson Natural Language Understanding*, a Amazon, com o *Amazon Comprehend*, e a Microsoft, com o *Azure Text Analytics*, de forma que, com uso das APIs fornecidas no formato de *Cloud Computing*, pode-se reduzir o tempo gasto do desenvolvedor e destinar esse recurso para o desenvolvimento dos negócios da empresa (MARKIEWICZ; ZHENG, 2020, p. 19).

Quando uma empresa estiver na dúvida se constrói inteiramente uma solução de PLN ou utiliza APIs, basta avaliar se a IA é um componente central para o ambiente de negócios, e essa avaliação pode direcionar se o nível técnico de conhecimento está envolvido na construção da aplicação da empresa. Como, por exemplo, uma empresa de e-commerce buscando inteligência adicional no suporte aos clientes, neste caso seria mais apropriado começar pelo uso de APIs, visto que o suporte ao cliente não é o *core* do seu negócio.

## 2 Referencial Teórico

O PLN é uma área de pesquisa que busca criar modelos computacionais que realizam tarefas específicas com linguagem natural, como por exemplo, a interpretação de textos e o processamento de voz. Essa área de pesquisa pode ser dividida em duas subáreas, sendo que a primeira busca lidar com os problemas que podem ser resolvidos por meio da análise de linguagem em nível sintático. São exemplos, problemas que esta primeira subárea busca lidar na revisão de texto, resumo automático da síntese de fala (baseada em texto), reconhecimento óptico de caracteres, extração de informações de um texto, síntese de fala (baseada em um texto) e sistemas simples de perguntas e respostas (FLASIŃSKI, 2016, p. 229).

A segunda subárea se dedica à resolução de problemas por meio da análise da linguagem em nível semântico. Os problemas que são resolvidos por essa subárea são exemplificados pela compreensão de fala e de texto, tradução

automática de texto de uma linguagem natural para outra linguagem natural, e criação de sistemas de comunicação verbal entre humanos e máquinas. Essa divisão em duas subáreas foi proposta porque hoje em dia apenas problemas pertencentes ao segundo grupo são desafiadores para IA.

Abordagens estatísticas para análise de linguagem começaram a ser desenvolvidas no final do século XX e têm como base o uso do *corpus* de texto, que são grandes conjuntos de textos referenciais para um determinado idioma. O *corpus* de texto é utilizado durante uma análise de texto com o auxílio de modelos para determinar características estatísticas do texto, como por exemplo, possíveis contextos em que uma determinada palavra aparece ou ainda possíveis usos de frases no texto.

Outra abordagem está relacionada à utilização de modelos gramaticais estocásticos e autômatos estocásticos, modelo esse que é equivalente ao modelo da cadeia de *Markov*<sup>5</sup> que também é utilizado em métodos avançados de PLN (FLASIŃSKI, 2016, p. 229).

Nas duas abordagens mencionadas acima, uma sintaxe pode ser utilizada como um ponto de partida para a análise da linguagem. Esse comportamento muitas vezes não é suficiente nos casos em que é necessário a compreensão dos conceitos envolvidos. Nesse caso, um sistema de IA que lide com esse tipo de problema precisaria ter um conhecimento adicional na forma de um modelo mundial. Essa questão pode ser resolvida definindo uma ontologia, sendo as redes semânticas uma das formas mais populares de defini-las.

A análise semântica computadorizada lida com os problemas complexos da linguagem falada, que tem como objetivo a comunicação entre humanos ou de humano e computador. Nesse contexto, muitos aspectos não verbais da comunicação expressos em linguagens não verbais, como a expressão facial e a linguagem corporal, podem conter pontos essenciais para compreensão. A quantidade de stress envolvido e a entonação, por exemplo, podem levar a uma mesma frase ser interpretada de maneiras muito diferentes (FLASIŃSKI, 2016, p. 230).

Para expressar uma mensagem com um sentido específico, o remetente deve passar o sentido apropriado utilizando características não verbais que revelam a sua intenção ao destinatário. No entanto, a compreensão apropriada da mensagem está relacionada a inteligência social. Esse tipo de inteligência é muito difícil de ser incorporado em um sistema de IA.

Como é possível observar, o PLN é uma área com muitas questões e abordagens envolvidas que podem ser consideradas como uma área bem desenvolvida da IA. Os exemplos de simulação de diálogos naturais por máquinas como *Chatbots* e o programa *Eliza*, de *Weizenbaum*, podem ser considerados exemplos de sucesso nessa área (FLASIŃSKI, 2016, p. 230).

---

<sup>5</sup> Trata-se de uma forma relativamente simples de modelar estatisticamente processos aleatórios.

Para melhor compreender a PLN, é necessário conhecer os principais contextos para realizar a análise e extração de dados em textos, onde se destacam entidades, relacionamentos, conceitos, sentimentos e emoções (MARKIEWICZ; ZHENG, 2020, p. 14):

- Entidades - o uso mais habitual dos contextos para PLN, sendo exemplos de entidades: pessoas, lugares, organizações e coisas contidas no texto;
- Relações - com o uso do PLN, é possível identificar relacionamentos entre as diversas entidades e igualmente informar o tipo de relação entre elas;
- Conceitos - uma das mais interessantes funções da PLN é a extração dos conceitos gerais do texto, que podem não ser aparentes no seu *corpus*.
- Palavras-chave - com uso da PLN é possível identificar quais são as palavras-chave, as mais importantes ou relevantes em um texto. Com esse recurso, é possível criar uma base de palavras-chave relacionadas ao *corpus*, que são importantes do ponto de vista comercial;
- Papéis semânticos - são formados pelos sujeitos e dizem respeito às ações e aos objetos com os quais agem no texto. Com PLN, é possível analisar as sentenças e extrair essas funções semânticas para atingir diversos fins comerciais, como, por exemplo, receber uma notificação sempre que uma determinada empresa lançar novos produtos;
- Categorias - descrevem sobre o que trata o conteúdo analisado. Neste contexto, o PLN pode realizar a análise e determinar a hierarquia taxonômica do conteúdo analisado, fornecendo, por exemplo as categorias de uso de um aplicativo, que, de acordo com o conteúdo, podem ser esportes, finanças, viagens, computação, entre outras. Pode-se destacar como uma possível utilização da categorização PLN para apresentar um conjunto de anúncios relevantes de acordo com o conteúdo apresentado em um site.
- Emoção - a detecção de emoções utilizando PLN pode ser um recurso extremamente valioso, pois com ela é permitido analisar se foi transmitida emoção de raiva, medo, alegria ou tristeza. Com tal recurso, é possível compreender de forma contínua e automática o sentimento relacionado a um determinado produto.

Sendo assim, foram apresentados alguns contextos para a construção de aplicações que utilizam a PLN.

### **3 Método**

Foi utilizada a metodologia de pesquisa bibliográfica para que fosse possível apresentar os seus principais conceitos, princípios arquitetônicos e 'estado da arte' das tecnologias envolvidas com a área de PLN e a criação de *Chatbots*. Sobre as perspectivas futuras será apresentado um panorama da evolução dos *Chatbots* presente na literatura, para situar o leitor em que nível estamos e qual o futuro dessa tecnologia. Além disso, é apresentado o



funcionamento de serviços de PLN oferecidos para criação de *Chatbots*, ofertados pelos fornecedores de *Cloud Computing*.

## 4 Resultados e Discussão

A seguir foram apresentados os resultados e discussões, onde foi contextualizado os desafios da área de PLN e apresentada a área de *Chatbots* através de uma visão geral, ascensão, implementação em ambiente de *Cloud Computing*, interface e tendências futuras para a área.

### 4.1 Desafios

Nem sempre o PLN é a solução para todos os problemas, mesmo tratando-se de uma tecnologia robusta. Neste contexto, existem duas principais áreas que se mostram como um desafio para a PLN. A primeira reside no fato de que a PLN funciona melhor com grandes conjuntos de dados e, conseqüentemente, quanto maior o número de dados, melhor é a precisão, muito embora o conjunto de dados necessários é dependente do tamanho real da aplicação, visto que, mais dados significam melhores resultados (MARKIEWICZ; ZHENG, 2020, p. 20; OLIVEIRA et al., 2021).

O segundo reside em considerar PLN uma bala de prata, pois quando se estuda PLN é comum visualizá-la como solução para os principais problemas, e quando realiza testes iniciais, os resultados tendem a ser muito precisos, pois a tendência é se utilizar de um corpo de texto simples em um primeiro momento, mas a linguagem humana tem muitas nuances. Conceitos como sarcasmo, gírias, jargões e humor são extremamente difíceis de serem interpretados, sem adicionarmos ao pacote as ambigüidades e os erros de ortografia ou gramática.

Nesses casos, até que a tecnologia consiga alcançar um alto grau de precisão, a melhor forma de lidar é filtrando o conteúdo que será processado com PLN. Enquanto não existe uma solução tecnológica, filtrar o conteúdo e identificar o problema com antecedência é a melhor estratégia a ser utilizada (MARKIEWICZ; ZHENG, 2020, p. 21; PASQUALOTTI; VIEIRA, 2008).

### 4.2 Chatbots

Um *Chatbot* é um serviço onde dados são expostos por meio de uma interface de PLN, representando uma excelente forma de exposição de informações para uma empresa, sendo fundamental compreender que, como interface, o *Chatbot* é tão bom como qualquer outro serviço de dados subjacentes. No entanto, antes da empresa investir na criação de um serviço de *Chatbot*, deve-se verificar a solidez dos dados (MARKIEWICZ; ZHENG, 2020, p. 24; ORSINI; FREITAS; COSTA; 2020).

Quando bem construídos, os *Chatbots* podem auxiliar a companhia a cortar custos e até estabelecer fluxos de receita adicionais, como exemplo, um agente virtual de suporte ao cliente pode reduzir o número de funcionários e dimensionar exponencialmente e com capacidade de suporte em tempo real. Outro exemplo reside em *Chatbot* de comércio conversacional, que dá ao seu

negócio um canal totalmente novo no qual a empresa pode envolver seus clientes por meio de plataforma de mensagens.

#### **4.2.1 A ascensão dos Chatbots**

Os *Chatbots* existem há mais de vinte anos, no entanto, eles estavam restritos a respostas como sim ou não, mas hoje apresenta-se um cenário totalmente diferente. Existem algumas razões pelas quais os cenários tenham tido grades mudanças, em especial o empreendedorismo (MARKIEWICZ; ZHENG, 2020, p.24; SILVA; LIMA, 2007):

- O aumento do recurso de PLN, em especial os disponíveis em serviços de *Cloud Computing*;
- A proliferação de serviços de mensagens populares, como o *Slack* e o *Facebook Messenger*;
- O aumento na criação de interfaces de linguagem natural.

#### **4.2.2 Chatbots em ambiente de Cloud Computing**

A disponibilização de recursos de PLN em formato de serviços de *Cloud Computing* tem sido a força motriz na ascensão das plataformas de *Chatbots*.

Serviços como classificadores de texto e extratores de entidades potencializaram algumas das principais funcionalidades dentro de um *Chatbot*. Técnicas de PLN não são recentes, mas ainda tem sido um problema utilizá-las fora do ambiente de pesquisa, de modo que uma alternativa para realizar a implementação desse tipo de técnica reside em soluções de código aberto, mas, com a chegada das APIs em ambiente de *Cloud Computing*, tornou-se muito mais simples para as empresas usarem essa tecnologia (MARKIEWICZ; ZHENG, 2020, p. 25).

#### **4.2.3 Interface de língua natural**

Os aplicativos de mensagens tomaram um espaço muito importante dentro do seguimento dos aplicativos móveis, ultrapassando as redes sociais em número de usuários ativos. Neste contexto, as empresas procuram maneiras de alcançar os usuários por meio desse importante canal de comunicação, gerando uma grande quantidade de dados contextuais nas mensagens onde são feitos questionamentos sobre lojas e buscas sobre mercadorias para compras, por exemplo (MARKIEWICZ; ZHENG, 2020, p. 25).

As empresas procuram incorporar os *Chatbots* no cotidiano dos usuários como uma plataforma de mensagens para obter respostas a dúvidas ou até mesmo auxiliar em diversas tarefas. Esse novo canal de comunicação muda a forma de interação homem-máquina, antes constituída pelo pressionar de teclas e cliques em botões do mouse. Assim, é extremamente interessante poder substituir essa forma da fala e linguagem natural, sendo esse um dos mais antigos objetivos da interação homem-máquina. Os mecanismos de busca permitem o uso de linguagem natural para que o usuário realize pesquisa na Web, mas atualmente a proliferação de dispositivos *Amazon Echo* atraiu desenvolvedores com a ideia de controlá-los para interagir com eletrodomésticos

por meio de agentes inteligentes (MARKIEWICZ; ZHENG, 2020, p. 25; FERRARETTO; KISCHINHEVSKY, 2010).

#### 4.2.4 Próxima geração de assistentes

Os exemplos de *Chatbots* que vemos ainda agregam pouco valor ao usuário final se comparados a sites ou aplicativos existentes. Mas existem exemplos poderosos de *Chatbots* de última geração que podem funcionar, auxiliando a empresa a gerar receitas adicionais ou economizar custos (NICHOL, 2018).

O caminho para os *Chatbots* é se transformarem em assistentes de IA que tornem possível aos clientes expressar aquilo que desejam em seus próprios termos, sem intervenção humana. Esses assistentes podem ser aplicações de atendimento ao cliente, e, para realizar operações em uma organização, a IA entende os clientes, no seu contexto e pode ser proativa, automatizando e agilizando muitas tarefas repetitivas.

Podemos classificar tais assistentes em cinco níveis, sempre caminhando em direção a uma organização autônoma. É possível visualizar cinco níveis de inteligência de assistentes de IA, são eles (NICHOL, 2018):

- Nível 1, os assistentes desse nível são conhecidos como assistentes de notificação, responsáveis por apresentar notificações de mensagens no *WhatsApp*, por exemplo, e outras notificações simples do telefone.
- Nível 2, os chamados assistentes de FAQ que são os exemplos mais comuns de assistentes, que permitem que o usuário faça uma pergunta simples e obtenha sua resposta de maneira simples, implementado um sistema mais ágil para as antigas páginas de FAQ.
- Nível 3, conhecidos como assistentes contextuais, onde o contexto é fundamental, o que torna o quê e como o usuário diz essenciais. Considerando o contexto, também significa ser capaz de entender e responder por entradas inesperadas.
- Nível 4, conhecidos como assistentes personalizados, são assistentes que aprendem quando é um bom momento para entrar em contato com base no contexto e sendo capaz de lembrar as preferências de cada um dos usuários, fornecendo uma interface personalizada e definitiva.
- Nível 5, são os assistentes de uma organização autônoma, que consistem em um grupo de assistentes que conhecem o cliente de maneira personalizada e executam grande parte das operações da empresa, desde a geração de leads, operações de RH, vendas ou finanças. Este tipo de assistente representa um grande salto e ainda levará alguns anos para se tornar realidade.

A grande maioria dos assistentes são os de FAQ, que parecem assistentes simples e em grande maioria de código aberto, sendo que a comunidade trabalha ativamente para elevá-los para uma assistente de nível 3 ou superior. A chave para isso reside nos contextos, que sem eles muitas situações não funcionam.

Um exemplo que ressalta tal situação no mercado é o fato de que apenas 2% dos 50 milhões de consumidores que possuem a *Alexa*, da *Amazon*, deixam habilitados seus dispositivos para realizar compras. Isso ocorre, pois, uma compra envolve muito do cenário e, se realizada muito rapidamente, pode gerar problemas, deixando claro que a interpretação de contexto ainda é insuficiente em relação às aplicações web e aplicativos.

Assistentes de nível 3 próximos, como vimos com o *Google Duplex*, que demonstrou como seria esse nível de assistente em um caso de uso específico. Neste assistente, o *Google* teve que inserir muita inteligência para fazê-lo funcionar em apenas um caso de uso e seria necessário muito mais inteligência para inserir um novo caso de uso, apresentando um produto de nicho (NICHOL, 2018; CORTESE, 2018).

Os seres humanos são capazes de generalizar e aplicar esse conteúdo em outros casos de uso e em diferentes contextos. Os assistentes de IA devem conseguir o mesmo resultado, mesmo que não estejam treinados especificamente, caso contrário eles não são escaláveis. A generalização implica que o assistente possa lidar com conversas de diferentes casos de uso semelhantes sem estarem treinados explicitamente com todas as possibilidades para os diferentes casos de uso (NICHOL, 2018).

## 5 Considerações finais

Em alguns anos veremos a evolução das tecnologias ligadas à PLN e conseqüentemente visualizaremos novas possibilidades e técnicas para construção de *Chatbots*. Fortes investimentos devem acontecer para que a ocorram treinamentos de novos modelos que ampliariam a capacidade da PLN e da subárea de *Chatbots*, onde a realização desta tarefa na forma de atendimentos poderia ser ampliada e melhorada.

Além disso, é pertinente refletir sobre os desafios enfrentados no PLN e ampliar as pesquisas com o objetivo de mitigar tais desafios. Outra questão que merece destaque é a evolução dos *Chatbots* onde o desenvolvimento da subárea de PLN contribuirá, ampliando e criando possibilidades para obter os benefícios econômicos da evolução dos *Chatbots* e possivelmente visualizar uma real evolução dos níveis de assistente.

Como trabalhos futuros acreditamos que o comparativo das técnicas de construção de *Chatbots* em especial em ambientes de *Cloud Computing*, poderá auxiliar os interessados em implementar esta tendência tecnológica para melhorar o atendimento ao cliente.

## Referências

CORTESE, João. Interação, indistinguibilidade e alteridade na Inteligência Artificial. **Teccogs: Revista Digital de Tecnologias Cognitivas**, n. 17, 2018.

FERRARETTO, L. A.; KISCHINHEVSKY, M. Rádio e Convergência: uma abordagem pela

economia política da comunicação. In: Compós 2010, 19, 2010, Rio de Janeiro. **Anais [...] Brasília**: Associação Nacional dos Programas de Pós-Graduação em Comunicação. Disponível em: <[http://www.compos.org.br/data/biblioteca\\_1495.pdf](http://www.compos.org.br/data/biblioteca_1495.pdf)>, acesso em 17 fev. 2019.

FLASIŃSKI, Mariusz. **Introduction to Artificial Intelligence**. Cham: Springer, 2016.

GDOURA, Khalil. **Is Neuro Linguistic Programming (NLP) Pseudoscience?** 2020. Disponível em: <<https://medium.com/illumination/curated/is-neuro-linguistic-programming-nlp-pseudoscience-7d02f166991d>>. Acesso em: 26 ago. 2022.

JOHRI, Prashant *et al.* Natural Language Processing: History, Evolution, Application, and Future Work. In: 3rd INTERNATIONAL CONFERENCE ON COMPUTING INFORMATICS AND NETWORKS. LECTURE NOTES IN NETWORKS AND SYSTEMS, 3., 2021. **Proceedings**. Singapore: Springer, 2021. v. 167, p. 365-375.

MARKIEWICZ, Tom; ZHENG, Josh. **Getting Started with Artificial Intelligence**: A Practical Guide to Building Enterprise Applications. 2. ed. Sebastopol, Ca: O'Reilly, 2020. 85 p.

MORGAN, Jasmine. **How Can Machine Learning Help Computers Understand Us?** 2018. Disponível em: <<https://www.bbvaopenmind.com/en/technology/artificial-intelligence/how-can-machine-learning-help-computers-understand-us/>>. Acesso em: 10 abr. 2021.

NICHOL, Alan. **The next generation of AI assistants in enterprise**. 2018. Disponível em: <<https://www.oreilly.com/radar/the-next-generation-of-ai-assistants-in-enterprise/>>. Acesso em: 24 fev. 2022.

OLIVEIRA, Bárbara Stéphanie Neves *et al.* Processamento de Linguagem Natural via Aprendizagem Profunda. **Jornada de Atualização em Informática 2022**, [S.L.], p. 8-54, 31 jul. 2022. SBC. <http://dx.doi.org/10.5753/sbc.9399.1.1>.

ORSINI, Adriana Goulart de Sena; FREITAS, Sérgio Henrique Zandonas; COSTA, Fabrício Veiga (coord.). Formas de Solução de Conflitos e Tecnologia. XI Congresso RECAJ-UFMG. Belo Horizonte: UFMG, 2020. 55 SALLES, Bruno Makowiecky. Jurisdição e Inteligência Artificial. **Revista da Escola Judiciária do Piauí**, Teresina, v. 2, n. 2, pp. 70-95, 2020.

PASQUALOTTI, Paulo Roberto; VIEIRA, Renata. WordnetAffectBR: uma base lexical de palavras de emoções para a língua portuguesa. **Renote**, Novo Hamburgo, Brasil, v. 6, n. 1, p. 1-10, 30 jun. 2008. Universidade Federal do Rio Grande do Sul. <http://dx.doi.org/10.22456/1679-1916.14693>.

SILVA, Renato Rocha; LIMA, Sérgio Muinhos Barroso. Consultas em bancos de dados utilizando linguagem natural. **Revista Eletrônica da Faculdade Metodista Granbery**, Juiz de Fora, v. 7, n. 2, ago/dez. 2007. Disponível em: <<http://re.granbery.edu.br/artigos/MjQ0.pdf>>. Acesso em: 30 ago. 2013.