

## CONTEXTO

“Tempo é dinheiro” (FRANKLIN, 1748) famosa frase dita por Benjamin Franklin na metade do século 18 ainda ecoa na cabeça de muitos seres humanos.

Em busca de mais tempo as pessoas procuram então otimizar suas tarefas. Uma forma de otimizar as tarefas é a criação de automatizações. As automatizações buscam por uma melhor produtividade, redução de custos e maior tempo livre para se concentrar em outras tarefas que não podem ser automatizadas. Para que isso seja possível, são integrados mecanismos que possam realizar as mesmas tarefas de maneira semelhante ou até mesmo com uma qualidade do melhor que do que o ser humano possa realizar.

Um grande exemplo disso são as assistentes virtuais, como por exemplo a Alexa da Amazon, a Siri da Apple e o Google Home do Google. Esses assistentes virtuais são capazes de realizar diversas tarefas através de um simples comando de voz. Essa interação entre seres humanos e máquinas está cada vez mais presente nos sistemas, mas para que isso seja possível, é utilizado uma ferramenta chamada PLN (Processamento de Linguagem Natural).

Johnson (2021) define o PLN sendo um ramo dentro da Inteligência Artificial responsável em fazer com que as máquinas possam compreender a linguagem dos seres humanos, ou seja, podemos dizer que o PLN funciona como um tradutor, permitindo assim que as tecnologias possam entender seus usuários, mesmo eles utilizando a linguagem natural.

O PLN também está presente em outras plataformas além das assistentes virtuais. Por exemplo, ele auxilia em sites de busca realizando interpretações entre o que o usuário digita com conteúdos de sites que poderão ser exibidos, estão presentes também no auto-completar em plataformas de busca onde sugestões automáticas são exibidas na tela no momento em que o usuário está digitando, e também em chatbots, que são utilizados por empresas para se comunicar com seus clientes, são robôs que funcionam dentro de aplicativos de mensagens que interpretam e respondem as interações de clientes com as empresas (TAKE BLIP, 2019).

Em todo processo de criação de um novo sistema, é necessário levantar os requisitos que esse sistema irá possuir. Em engenharia de requisitos, a etapa responsável para o levantamento dessas informações é a elicitacão.

Para Thayer (1997), a elicitação de requisitos é o processo em que os clientes e usuários são questionados pela equipe de desenvolvimento a falarem o que espera como funcionalidades no sistema que será desenvolvido.

Nessa etapa de elicitação serão definidas as exigências, os recursos, os objetivos e as utilidades que o sistema deve cumprir.

Segundo Sommerville(2011, pág. 57):

Os requisitos de um sistema são as descrições do que o sistema deve fazer, os serviços que oferece e as restrições de seu funcionamento. Esses requisitos refletem as necessidades dos clientes para um sistema que serve a uma funcionalidade determinada, como controlar um dispositivo, colocar um pedido ou encontrar informações. O processo de descobrir, analisar, documentar e verificar esses serviços e restrições é chamado engenharia de requisitos.

Uma das técnicas utilizadas no desenvolvimento de software ágil é a criação de Histórias de Usuário (User Stories). Através delas, o usuário utiliza de uma abordagem de escrever sobre os requisitos, tudo isso através de uma ou duas frases escritas através da perspectiva de quem deseja o recurso/funcionalidade.

Para Cohn (2009, pág. 4), “uma história de usuário descreve a funcionalidade que será valiosa para um usuário ou comprador de um sistema ou software.”

Rehkopf (2020) define histórias de usuário como “uma explicação informal e geral sobre um recurso de software escrita a partir da perspectiva do usuário final. Seu objetivo é articular como um recurso de software pode gerar valor para o cliente.”

As técnicas de PLN também podem oferecer diversas vantagens para melhorar a qualidade das histórias de usuário. Segundo Raharjana, Siahaan e Fatichah (2021):

As técnicas de processamento de linguagem natural (PNL) oferecem vantagens potenciais para melhorar a qualidade das histórias de usuários. O PNL pode ser usado para analisar, extrair ou analisar os dados da história do usuário. Tem sido amplamente utilizado para ajudar no domínio da engenharia de software (por exemplo, gerenciamento de requisitos de software, extração de atores e ações no documento de requisitos, recurso de software extração, teste de software, etc.).

## **PROBLEMA**

Johnson (2021) divide o PLN em cinco componentes de processamento: análise morfológica e lexical, análise sintática, análise semântica, integração de discurso e análise pragmática.

Análise morfológica e lexical, divide um texto em parágrafos, palavras e frases identificando a estrutura das palavras. Análise sintática, é responsável pela ordenação adequada das palavras, o que pode afetar seu significado. Isso envolve a análise das palavras em uma frase, seguindo a estrutura gramatical da frase. Análise semântica é uma estrutura criada pelo analisador sintático que atribui significados. Essa análise concentra-se apenas no significado literal de palavras, frases e sentenças. Isso apenas abstrai o significado do dicionário ou o significado real do contexto dado. Integração de discurso foca na análise de contexto da frase. Ou seja, se fossemos analisar a palavra “aquilo” na frase “Não posso pegar aquilo”, o significado dependerá do discurso anterior. Por fim, a análise pragmática lida com o conteúdo comunicativo e seu efeito na interpretação. Ou seja, abstrai ou deriva o uso significativo da linguagem. Nessa análise, o foco principal é sempre no que foi dito e é reinterpretado sobre o que se quer dizer (Johnson, 2021).

Atualmente existem inúmeras soluções utilizadas para o PLN, Parker(2019) em seu artigo cita 12 ferramentas open source em diversas linguagens de programação, como por exemplo Python, Node e Java.

Dentre as soluções citadas por Parker (2019), destaca-se a Natural Language Toolkit (NLTK) em Python, por ser a solução com mais recursos disponíveis, capaz de implementar todos os componentes de PLN e oferece suporte a vários idiomas. Outra solução que se destaca é a OpenNLP em Java. É hospedada pela Apache Foundation, ou seja, é fácil integrá-la com outros serviços da Apache. Assim como a NLTK, oferece suporte a vários idiomas e cobre todos os componentes de PLN.

Cohn (2009) comenta que ao definir os requisitos de software a comunicação pode ser um adversidade, pois aqueles que desejam um novo software devem se comunicar com quem irá desenvolvê-lo. Caso uma das partes dominem a comunicação, o projeto poderá ter problemas. Se a equipe de negócios dominar, ela exigirá funcionalidades e datas sem se preocupar se os desenvolvedores poderão atender a demanda ou até mesmo saber se eles compreenderam o que foi proposto ou não. Caso a equipe técnica domine a comunicação, termos técnicos podem substituir a linguagem de negócio perdendo a oportunidade de aprender mais sobre o negócio do sistema que será desenvolvido.

Se tratando de histórias de usuário, Heath (2020) afirma que muitas vezes são genéricas, vagas na coleta e análise de requisitos. Segundo o autor, ele afirma que as maiores dificuldades encontradas são referentes a escopo, priorização e

classificação das histórias no backlog, imprecisão sobre o que é problema e o que é solução, e por fim, histórias que escondem seu real valor.

Sendo assim, como selecionar a solução mais adequada para auxiliar o PLN dentro do contexto de histórias de usuário?

Neste sentido, o objetivo central de desse trabalho é realizar um estudo sistemático de comparação entre pequenas soluções utilizando PLN para criação de histórias de usuário para os idiomas português e inglês, e por fim, serão avaliadas a qualidade das histórias de usuário criadas.

## REFERÊNCIAS BIBLIOGRÁFICAS

BARKER. D. **"12 open source tools for natural language processing"**, 2019.

Disponível em:

<https://opensource.com/article/19/3/natural-language-processing-tools> Acesso em 11 dez. 2021

COHN, M. **User stories applied for agile software development**, 13. ed.

Crawfordsville, Indiana. 2009. 263 p.

FRANKLIN. B. **"Advice to a Young Tradesman"**, 1748. Disponível em:

<https://founders.archives.gov/documents/Franklin/01-03-02-0130> Acesso em 10 dez. 2021

HEATH, F. **The trouble with user stories**. 2020, DZone. Disponível em:

<https://dzone.com/articles/the-trouble-with-user-stories-1> . Acesso em 04 dez. 2021

JOHNSON. D. **"Natural Language Processing Tutorial: What is NLP? Examples"**,

2021 <https://www.guru99.com/nlp-tutorial.html> . Acesso em 09 dez. 2021

RAHARJANA, I. K, SIAHAAN. D e FATICHAH. C, **"User Stories and Natural**

**Language Processing: A Systematic Literature Review,"** in IEEE Access, vol. 9, pp. 53811-53826, 2021, doi: 10.1109/ACCESS.2021.3070606.

REHKOPF, M. **“Histórias de usuários com exemplos e um template”**, 2020.

Disponível em: <https://www.atlassian.com/br/agile/project-management/user-stories> .

Acesso em 09 dez. 2021

SOMMERVILLE, I. . **Engenharia de software**, 9. ed. Pearson. 2011. 529 p.

TAKE BLIP, **Tudo sobre NLP: o que é processamento de linguagem natural e seus desafios na Inteligência Artificial**, 2019. Disponível em:

<https://www.take.net/blog/tecnologia/nlp-processamento-linguagem-natural/> . Acesso em 01 dez. 2021.

THAYER, R. H. e DORFMAN, M.; **“Introduction to Tutorial Software Requirements Engineering”** in Software Requirements Engineering, IEEE-CS Press, Second Edition, 1997, p.p. 1-2.