

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ  
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIA EM SAÚDE**

**JOSÉ FERNANDO ANSELMO GUIMARÃES**

**ESTRUTURAÇÃO DE LAUDOS DE IMAGENS MÉDICAS  
ATRAVÉS DE PROCESSAMENTO DE LINGUAGEM NATURAL**

**CURITIBA**

**2014**

**JOSÉ FERNANDO ANSELMO GUIMARÃES**

**ESTRUTURAÇÃO DE LAUDOS DE IMAGENS MÉDICAS  
ATRAVÉS DE PROCESSAMENTO DE LINGUAGEM NATURAL**

Dissertação de mestrado apresentado ao Programa de Pós-Graduação em Tecnologia em Saúde da Escola Politécnica, da Pontifícia Universidade Católica do Paraná, como requisito parcial à obtenção do título de mestre.

Orientador: Prof. Dr. Cláudia M<sup>a</sup> Cabral Moro Barra  
Coorientador: Prof. Dr. Luiz Roberto Aguiar

**CURITIBA**

**2014**



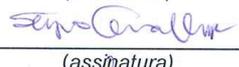
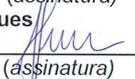
Pontifícia Universidade Católica do Paraná  
Escola Politécnica  
Programa de Pós-Graduação em Tecnologia em Saúde

**ATA DE DEFESA DE DISSERTAÇÃO DE MESTRADO  
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIA EM SAÚDE**

**DEFESA DE DISSERTAÇÃO Nº 197**

**ÁREA DE CONCENTRAÇÃO: TECNOLOGIA EM SAÚDE**

Aos 14 dias do mês de agosto de 2014, na sala hum de mestrado, realizou-se a sessão pública de Defesa da Dissertação: "Estruturação de Laudos de Imagens Médicas Através de Processamento de Linguagem Natural", apresentada pelo aluno José Fernando Anselmo Guimarães, sob orientação do Profª. Drª. Claudia Maria Cabral Moro Barra e coorientação do Prof. Dr. Luiz Roberto Aguiar, como requisito parcial para a obtenção do título de Mestre em Tecnologia em Saúde, perante uma Banca Examinadora composta pelos seguintes membros:

Profª. Drª. Claudia Maria Cabral Moro Barra, PUCPR (Orientador e presidente)	 (assinatura)	<u>APROVADA</u> (Aprov/Reprov.)
Prof. Dr. Sérgio Ossamu Ioshii, PUCPR (Examinador)	 (assinatura)	<u>APROVADO</u> (Aprov/Reprov.)
Prof. Dr. Paulo Mazzoncini de Azevedo Marques USP (Examinador)	 (assinatura)	<u>APROVADO</u> (Aprov/Reprov.)

Início: 14:40 Término: 16:40

Conforme as normas regimentais do PPGTS e da PUCPR, o trabalho apresentado foi considerado APROVADO (aprovado/reprovado), segundo avaliação da maioria dos membros desta Banca Examinadora.

Observações: NÃO HA

O aluno está ciente que a homologação deste resultado está condicionada: (I) ao cumprimento integral das solicitações da Banca Examinadora, que determina um prazo de 60 dias para o cumprimento dos requisitos; (II) entrega da dissertação em conformidade com as normas especificadas no Regulamento do PPGTS/PUCPR; (III) entrega da documentação necessária para elaboração do Diploma.

ALUNO: José Fernando Anselmo Guimarães   
(assinatura)

  
Profª. Drª. Claudia Maria Cabral Moro Barra,  
Coordenadora do PPGTS PUCPR

*o Prof. Dr. Paulo Mazzoncini autorizou o*

*Prof. Luiz Roberto Aguiar assinar por ele.*

## **AGRADECIMENTOS**

As equipes de desenvolvimento e de teste especialmente as enfermeiras que se ofereceram para fazer o trabalho. Ao professor Aguiar, meu co-orientador que deu importantes direcionamentos no início do trabalho e não só liderou os testes finais com sua equipe, mas participou ativamente dos mesmos. A professora Cláudia, minha orientadora, que sempre presente, soube conciliar a gentileza e a firmeza, a pressa com a tranquilidade e o sonho com a realidade. Especialmente a minha grande amiga Déborah, que me incentivou a entrar no mestrado, participou do processo sempre com bons conselhos e observações, e que demonstrou mais uma vez ser ótima professora, profissional indiscutível e uma pessoa extraordinária.

É evidente que todas as ciências têm, em maior ou menor grau, uma relação com a natureza humana, e que por mais que qualquer delas pareça afastar-se, sempre retornam a ela em uma ou outra passagem.

(HUME, 1739)

## RESUMO

Pesquisadores e associações médicas têm defendido a estruturação de laudos radiológicos. Entre os objetivos estão facilitar tanto o entendimento por parte do profissional que solicitou o exame correspondente, como o acesso aos grandes volumes de laudos hoje presentes nos sistemas de informação em saúde. Entre as alternativas para obter os laudos estruturados está a utilização de Processamento de Linguagem Natural (PLN) sobre as bases dos laudos sem estruturação prévia. No entanto, há pouca disponibilidade de sistemas de PLN para a língua portuguesa, em especial os que identifiquem informações de forma abrangente e sem depender de um dicionário ou lista específica. Alguns trabalhos que utilizam regras morfológicas possibilitam a identificação de informações em narrativas clínicas de uma forma mais genérica, porém não foram aplicadas a laudos radiológicos. Também a utilização de sistemas de PLN diretamente por profissionais de saúde depende de suporte ou treinamento, pois esta não é uma área de conhecimento típica para os mesmos. **Objetivo:** Desenvolver um método e um protótipo de sistema que permitam a utilização de técnicas de PLN para a estruturação de laudos radiológicos de ressonância magnética de encéfalo que possam ser utilizados diretamente por profissionais de saúde. **Método:** Trata-se de uma pesquisa metodológica e exploratória com abordagem quantitativa. Foi utilizada a técnica de regras de classificação morfológica. Foi dividida em 4 fases: Desenvolvimento do protótipo, Criação e execução das regras, Utilização do protótipo por profissionais de saúde e a Proposição do método. **Resultados:** Método para criação das regras utilizando o protótipo, sendo que na sua utilização pela equipe de desenvolvimento foram criadas 182 regras e foi obtido precisão de 0,74, sensibilidade de 0,81 e *f-measure* de 0,77. Com a equipe de teste foi obtido precisão de 0,59, sensibilidade de 0,80 e *f-measure* de 0,68. **Conclusão:** A utilização da técnica de regras de classificação morfológica mostrou-se útil para os objetivos pretendidos, apesar de poucas pesquisas terem sido encontradas. A evolução do método utilizado anteriormente permitiu encontrar novos usos e identificar limites na utilização da mesma.

**Palavras-chave:** Processamento de linguagem natural; Recuperação de informação; Laudo; Radiológico.

## ABSTRACT

Researchers and medical associations have been defending the structuring of radiological reports. Two of its goals are facilitating the understanding by the medical physicist who requests the exam, and the access to the large volume of data in the health informatics systems. Among the alternatives to get radiological reports structured is applying a Natural Language Processing (NLP) over the nonstructured radiological reports data bases. However, there are few NLP systems available in Portuguese, especially the ones which are able to widely identify information without the need of a dictionary or specific list of terms. Some researches use morphological classification rules and enable the identification of information in clinical narratives in a more generic form, but it was not applied to radiological reports. Another issue is the need of support and training for health professionals so that they can use these PLN systems, as it is not a typical knowledge for them. **Objective:** To develop a method and a system prototype which allows the use of NLP to radiological reports of magnetic resonance electroencephalography that could be used directly by health professionals. **Method:** It is a methodological and exploratory research with quantitative approach. It has used the morphological classification rules technique. It was performed in 4 phases: Prototype development, rules creation and execution, prototype used by health professionals and method proposal. **Results:** The method for rules creation used the prototype, wherein the development team created 182 rules and the precision was 0,74, the sensibility was 0,81 and *f-measure* was 0,77. The precision of the test team was 0,59, the sensibility was 0,80 and the *f-measure* was 0,68. **Conclusion:** The use of the morphological classification rules shows to be useful to the intended goals, although few researches about this technique have been found. The evolution of the previous method allowed finding new uses and identifying limits of the technique.

**Keywords:** Natural Language Processing; Radiology; Report; Health System.

## LISTA DE FIGURAS

Figura 1 – Texto de exemplo para mostrar algoritmo .....	30
Figura 2 – Fluxograma do passo 3 .....	35
Figura 3 – Fluxograma do passo 4 .....	37
Figura 4 – Tela com a lista de arquivos .....	42
Figura 5 – Tela de cadastramento de regras .....	43
Figura 6 – Tela de escolha das regras para execução .....	43
Figura 7 – Tela de resultados .....	44
Figura 8 – Equipe de desenvolvimento – VP e FP por quantidade de elementos das regras .....	47
Figura 9 – Laudo de ressonância magnética antes do pré-processamento .....	49
Figura 10 – Laudo de ressonância magnética depois do pré-processamento .....	49
Figura 11 – Equipe de teste - VP e FP por quantidade de termos das regras .....	52

## LISTA DE QUADROS

Quadro 1 – Significado das etiquetas utilizadas .....	20
Quadro 2 – Fases e processos utilizados por Oliveira (2014) .....	22
Quadro 3 – Técnicas e objetivos dos artigos com a quantidade de laudos utilizados .....	26
Quadro 4 – Fases e etapas do trabalho .....	27
Quadro 5 – Quantidades de laudos utilizados em cada atividade .....	28
Quadro 6 – Utilização de regra em laudo radiológico .....	35
Quadro 7 – Etiquetas utilizadas e seus significados .....	36
Quadro 8 – Estruturação do laudo da Figura 9 .....	50

## LISTA DE TABELAS

Tabela 1 – Equipe de desenvolvimento - Quantidade de regras por grupo de 5 textos .....	46
Tabela 2 – Equipe de desenvolvimento - Quantidade de regras por número de termos, com a quantidade de achados no texto e quantidade de VP e FP .....	46
Tabela 3 – Equipe de desenvolvimento - Valores obtidos .....	47
Tabela 4 – Equipe de desenvolvimento - Média e desvio padrão de VP e FP por texto .....	47
Tabela 5 – Equipe de desenvolvimento - Resultados das melhores regras .....	48
Tabela 6 – Equipe de desenvolvimento - Resultados das piores regras .....	48
Tabela 7 – Equipe de Teste - Quantidade de regras por grupo de 5 textos .....	51
Tabela 8 – Equipe de Teste - Quantidade de regras por número de termos, com a quantidade de achados no texto e quantidade de VP e FP .....	51
Tabela 9 – Equipe de Teste - Valores obtidos .....	52
Tabela 10 – Equipe de Teste - Média e desvio padrão de VP e FP por texto .....	52
Tabela 11 – Equipe de Teste - Resultados das melhores regras .....	53
Tabela 12 – Equipe de Teste - Resultados das piores regras .....	53

## LISTA DE ABREVIATURAS E SIGLAS

AR	Achado Radiológico
CoGrOO	Corretor gramatical para (Apache Open Libre) Office
ESTEMED	Sistema Estruturação de Textos Médicos
FN	Falso Negativo
FP	Falso Positivo
PLN	Processamento de Linguagem Natural
SNOMED	<i>Systematized Nomenclature of Human Medicine</i>
TM	Termo Morfológico
UML	<i>Unified Modeling Language</i>
VP	Verdadeiro Positivo

## Sumário

1 INTRODUÇÃO .....	14
1.1 OBJETIVOS .....	17
1.1.1 Objetivo Geral .....	17
1.1.2 Objetivos Específicos .....	17
2 REFERENCIAL TEÓRICO .....	18
2.1 RECUPERAÇÃO DE INFORMAÇÃO DE TEXTOS MÉDICOS UTILIZANDO PLN .....	18
2.2 PLN EM LAUDOS RADIOLÓGICOS .....	23
3 MÉTODO .....	27
3.1 BASE DE DADOS .....	28
3.2 LOCAL DO ESTUDO .....	29
3.3 PARTICIPANTES .....	29
3.4 O DESENVOLVIMENTO DO PROTÓTIPO DE SISTEMA ESTEMED .....	30
3.5 CRIAÇÃO E EXECUÇÃO DAS REGRAS .....	33
3.5.1 Execução das regras .....	39
3.6 UTILIZAÇÃO DO ESTEMED PELA EQUIPE DE TESTE .....	39
3.7 PROPOSIÇÃO DO MÉTODO .....	40
4 RESULTADOS .....	41
4.1 O ESTEMED .....	41
4.2 RESULTADOS DA APLICAÇÃO DO ESTEMED .....	45
4.2.1 Utilização do sistema pela equipe de desenvolvimento .....	45
4.2.2 Utilização do ESTEMED por profissionais de saúde .....	50
4.3 MÉTODO PROPOSTO .....	53
5 DISCUSSÃO .....	57
5.1 O MÉTODO .....	57
5.2 O PROCESSO DA CRIAÇÃO DAS REGRAS E A UTILIZAÇÃO DO MÉTODO NO ESTEMED .....	60
5.3 TRABALHOS FUTUROS .....	62
6 CONCLUSÃO .....	64

## 1 INTRODUÇÃO

Pesquisadores têm enfatizado os benefícios da estruturação de textos de laudos radiológicos (Pool e Goergen, 2010; Reiner, Knight e Siegel, 2007; Kahn, Langlotz e Burnside, 2009). Os objetivos desta estruturação são: melhor comunicação entre o médico que solicita e o que realiza o exame; e maior facilidade na recuperação dos laudos para revisão do seu conteúdo, cobrança, auditoria, treinamento, pesquisa e ensino (Kahn, Langlotz e Burnside, 2009).

A estruturação do texto do laudo pode ser obtida pela própria disciplina do profissional durante a sua elaboração, escrevendo-os de uma forma pré-definida; ou utilizando um sistema de informação clínica que tenha a funcionalidade de estruturação para digitação (Kahn, Langlotz e Burnside, 2009).

Outra forma de se obter esta estruturação é utilizando técnicas de processamento em linguagem natural (PLN). Originalmente, o PLN era uma intersecção entre a inteligência artificial e a linguística, e com o tempo, todavia, convergiu com a área de recuperação de informação. Ele empresta técnicas de diversas áreas de conhecimento e torna possível realizar manipulações em textos em geral (Chapman e Cohen, 2009).

O PLN permite a manipulação também de textos médicos, como já foi feito por pesquisadores (Do *et al.*, 2012; Dublin *et al.*, 2013; Yetisgen-Yildiz *et al.*, 2013), que recuperam informações específicas, como fraturas em ossos, identificação de pneumonia e recomendações dos laudos. Para a estruturação de textos médicos estas técnicas devem ser adaptadas para não recuperarem apenas uma determinada informação, mas sim, todas as informações relevantes.

Desta forma, Serapião *et al.* (2013), Honorato *et al.* (2008), Gerstmair *et al.* (2012), Lacson *et al.* (2012) e Prasad e Ramakrishna (2010) aplicaram técnicas de PLN em textos com objetivo de estruturação, construção de um dicionário, mapeamento de informações não estruturadas e criação de sistemas de busca. Estas pesquisas utilizam técnicas de recuperação de informação extraíndo todas as informações consideradas relevantes.

Deve-se também considerar o entendimento do que seja um laudo estruturado. Encontra-se na literatura a observação de que os laudos radiológicos não são estruturados. Algum nível de estruturação é encontrada em diversos

sistemas utilizados na criação de laudos, pois os mesmos têm frases padrão pré-definidas para explicitar diversas situações de normalidade, e estas frases são muitas vezes alteradas para explicitar uma situação específica de anormalidade. No entanto, não é frequente uma definição mais específica do que seria uma estruturação e seu conceito, que varia entre diferentes autores. Prasad e Ramakrishna (2010) consideram a estruturação do laudo como a identificação de achados médicos, permitindo identificar a ocorrência, o lugar à que se refere e as dimensões, quando houverem. Já Kahn, Langlotz e Burnside (2009) afirmam que seria ideal que um laudo estruturado fosse dividido em seções significativas, ordenadas de forma consistente, contendo linguagem padronizada.

Achados radiológicos (AR) e termos morfológicos (TM) são alguns exemplos de informações (achados clínicos) identificadas para realizar uma estruturação parcial de laudos. AR refere-se a “o que”, como exemplo: “apresentam um aspecto quadrangular”, “estenose moderada” e “focos de hipossinal em t1 e hipersinal em t2”; e TM são o “onde”. São exemplos de TM: “ventrículos cerebrais supratentoriais”, “carótida interna esquerda” e “tálamo e região periaquedutal à esquerda”.

As técnicas utilizadas, individualmente ou combinadas, em PLN são basicamente de dois tipos (Climenson, 1963; Rezende, 2005):

1) Baseadas em estatística (Yetisgen-Yildiz *et al.*, 2013). Nesta técnica é realizada uma contagem de palavras em cada texto e no conjunto deles. São calculados diferentes índices para cada palavra, bem como para os agrupamentos de palavras. Assim é possível inferir a importância de uma palavra ou agrupamento.

2) Baseada em regras (Bui *et al.*, 2010; Percha *et al.*, 2012). São técnicas que identificam uma sequência de determinada característica do texto. Formam-se então regras que relacionam trechos de interesse do texto à sequência destas características. Quando estas regras são aplicadas ao texto e são encontradas sequências idênticas, então conclui-se que são do mesmo tipo de informação que os trechos originais correspondentes. As características, que podem ser utilizadas em conjunto, são classificações morfológicas, sintáticas ou partes da palavra propriamente dita, como raiz, sufixos e prefixos.

As técnicas de regras e estatística podem ser utilizadas em conjunto, e além delas, muitas vezes também é utilizado um dicionário de forma auxiliar (Friedman,

2000). Desta forma, cada palavra ou conjunto de palavras são identificados como uma ocorrência em um dicionário, a partir do qual se obtém informações relevantes, como a classificação morfológica, um sinônimo, um código referente a algum dicionário ou nomenclatura de termos da área médica (como o *Systematized Nomenclature of Human Medicine* (SNOMED)) e que tipo de informação está associada a esta palavra.

Os pesquisadores normalmente não citam as dificuldades na utilização das técnicas de PLN. Como as mesmas não fazem parte das atividades-fim da área médica, os profissionais desta área necessitariam de suporte de alguém especializado ou de treinamento. No caso de sistemas que utilizam dicionários, há a dificuldade de formação dos mesmos, que é manual e demanda um esforço significativo (Gerstmair *et al.*, 2012). A utilização de regras morfológicas, depende de um *corpus* anotado da área de conhecimento de cujos textos se está analisando, que é um *corpus* com etiquetas associadas a cada palavra que identificam características das mesmas, como classificação morfológica, classificação sintática e raiz da palavra. A formação de um *corpus* anotado também pode demandar um certo trabalho, muito inferior ao da criação de um dicionário, mas mesmo assim, deve ser considerado.

Nas pesquisas de Bulegon (2011) e Oliveira (2014) foram utilizadas regras com classificações morfológicas baseadas em um *corpus* anotado de resumos de alta. A criação das referidas regras foi manual ou parcialmente automatizada, o que representa uma dificuldade para um usuário sem conhecimento prévio.

Tem-se então como hipóteses que o trabalho necessário para criação de um *corpus* anotado com textos de laudos de ressonância magnética pode ser evitado utilizando-se o *corpus* anotado de resumos de alta, e que o trabalho de criação de regras de classificações morfológicas pode ser simplificado através do desenvolvimento de um sistema que seja suficientemente fácil de usar para profissionais de saúde. A confirmação das mesmas minimiza o problema de dificuldades dos profissionais de saúde na utilização de sistemas de PLN e o problema da grande quantidade de recursos necessários nas atividades correlatas.

## 1.1 OBJETIVOS

### 1.1.1 Objetivo Geral

Elaborar um método para estruturação de textos de laudos radiológicos de ressonância magnética, escritos no idioma português, que possa ser utilizado por profissionais de saúde, minimizando a necessidade de suporte ou treinamento.

### 1.1.2 Objetivos Específicos

Os objetivos específicos foram:

- Construir protótipo do sistema de Estruturação de Textos Médicos (ESTEMED) utilizando a técnica de regras de classificação morfológica e um corpus anotado de textos de resumos de alta.
- Utilizar o ESTEMED e verificar seus resultados.
- Verificar a facilidade na utilização do ESTEMED por profissionais de saúde.
- Especificar a sequência de passos para o profissional de saúde utilizar o ESTEMED.

## 2 REFERENCIAL TEÓRICO

Como o objetivo do trabalho está ligado à estruturação de laudos, que é um documento essencial em qualquer procedimento de imagem (Kahn, Langlotz e Burnside, 2009), e que a estruturação foi feita utilizando técnica de PLN, o primeiro item da revisão será recuperação de informação utilizando PNL em textos médicos.

A seguir são descritos os trabalhos de PLN em laudos radiológicos, incluindo uma descrição do que é o laudo radiológico, e que tipo de informação está sendo pesquisada nestes textos.

### 2.1 RECUPERAÇÃO DE INFORMAÇÃO DE TEXTOS MÉDICOS UTILIZANDO PLN

O PLN faz parte da área de inteligência artificial. É um conjunto de técnicas e métodos transformados em sistemas que processam os textos em linguagem natural com objetivos como categorização automática de um texto (Stamatatos, Fakotakis e Kokkinakis, 2005) extração de informações específicas como doenças ou diagnósticos (Bulegon, 2011; Childs *et al.*, 2009; Yang *et al.*, 2009), realização de processos de indexação, nos quais são encontradas palavras-chave do texto, relevantes para facilitar uma futura pesquisa do mesmo (Schulz *et al.*, 2002; Lacson *et al.*, 2012) e o *clustering* (Patterson e Hurdle, 2011), que utilizam técnicas de mineração de dados, para agrupar os textos de narrativas clínicas, de forma que o pesquisador pode analisar as semelhanças entre os agrupamentos de textos encontrados com esta técnica.

Também a estruturação dos textos médicos pode ser realizada através de técnicas de PLN. Com ela pode-se identificar informações relevantes e interessantes nos textos (Prasad e Ramakrishna, 2010).

Os métodos de PLN para recuperação de informação identificados na literatura podem ser divididos em dois grandes grupos, segundo Climenson (1963) e Rezende (2005), o baseado em regras e o baseado em estatísticas. Na área médica podemos encontrar recuperação de informações, como por exemplo, de medicações (Yang, 2010), de relações entre informações médicas (Dai *et al.*, 2012), de uso

excessivo de medicamentos (Salmasian e Freedberg, 2012) e de diagnósticos (Bulegon, 2011) que são baseados em regras; e de estado de comorbidades (Ambert e Cohen, 2009), de informações farmacológicas (Wang *et al.*, 2009) e de classificação de doenças (Li, Chen e Huang, 2011) que são baseados em estatísticas.

Nos métodos baseados em estatística cada texto é relacionado com todas as palavras relevantes que o compõe, excluindo palavras de pouco significado, como artigos e preposições, chamadas de *stop-words*, e é então realizada a contagem destas palavras. Com as estatísticas das ocorrências das palavras em cada texto e em todos os textos do conjunto são utilizadas técnicas matemáticas e estatísticas para avaliar se a palavra é importante ou não. Os mesmos cálculos também são realizados para grupos de palavras, chamados *n-grams* (Yetisgen-Yildiz *et al.*, 2013). Assim sendo, uma frase como “Após a administração de contraste observa-se ...” teria as *stop-words* eliminadas (“a” e “de”), e da mesma forma que cada palavra é contada, os agrupamentos dos diversos tamanhos também são. Se fosse arbitrado o *n-gram* máximo de 3 palavras, teríamos os seguintes *n-grams* para as palavras “Após”, “administração” e “contraste”: “Após administração”, “administração contraste” e “Após administração contraste”.

Nos métodos baseados em regras o processo é composto por duas fases principais: formação das regras e aplicação das regras nos textos.

As regras representam uma sequência de características das palavras do texto, como por exemplo, a classificação morfológica ou gramatical (Bulegon, 2011), a função gramatical ou sintática ou a palavra inteira ou partes dela, como a raiz (Percha *et al.*, 2012), o prefixo ou o sufixo.

Um exemplo de regras baseadas em partes de palavras pode ser encontrado em Percha *et al.* (2012) onde é procurado no texto as ocorrências de “*predominantly fatty*” e possíveis variações, com a regra “*predomin\* fat(ty)*”. Nesta regra está especificado para identificar qualquer palavra que inicie por “*predomin*” e que tenha em seguida a palavra “*fat*” ou “*fatty*”.

Nos métodos baseados em regras de classificação morfológica, num primeiro passo, determina-se a classe morfológica de todas as palavras de um texto, associando às palavras uma etiqueta de classificação morfológica. Depois, pode-se

estudar a sequência destas etiquetas que identificam uma determinada informação que queremos recuperar nos textos. Por exemplo, na pesquisa de Bulegon (2011), que utilizou resumos de alta, se for identificada uma informação que se deseja recuperar com a sequência de etiquetas morfológicas “pronome, substantivo, adjetivo”, quando esta sequência ocorrer novamente em algum outro texto, se presume que também seja uma ocorrência da mesma informação de interesse. Esta sequência de etiquetas morfológicas é chamada de regra.

No método de regras de classificações morfológicas utilizada por Bulegon (2011), devido à necessidade de se trabalhar com textos em português, foram utilizadas bibliotecas da ferramenta “Corretor gramatical para (Apache Open|Libre) Office” (CoGrOO), que atende bem a estas necessidades, para colocar as etiquetas morfológicas nas palavras.

Quadro 1 - Significado das etiquetas utilizadas.

[N_M_S]	Substantivo ( <i>Noum</i> ), masculino, singular.
[N_F_S]	Substantivo, feminino, singular.
[N_M_P]	Substantivo, masculino, plural.
[ADJ_M_S]	Adjetivo, masculino, singular.
[ADJ_F_S]	Adjetivo, feminino, singular.
[KC]	Conjunção Coordenada.
[PRP]	Preposição.
[V_PS_3S_IND_VFIN]	Verbo, perfeito simples, 3a pessoal do singular, indicativo, finito.
[PNT_ABS]	Ponto final.
[PNT_NSEP]	Vírgula.

Fonte: Bulegon, 2011.

O CoGrOO é uma ferramenta *open source* e foi desenvolvida em Java. Utilizando-se as bibliotecas do CoGrOO pode-se identificar para cada palavra no texto sua classificação, disponibilizando uma etiqueta correspondente, conforme Quadro 1.

No seu trabalho, Bulegon (2011) identifica diagnósticos médicos nos textos de resumos de alta pela sequência de classificação de três termos consecutivos. Por exemplo, “com angina pectoris” tem a sequência “[PRP] [N\_F\_S] [ADJ\_F\_P]”

(conforme etiquetas do Quadro 1). Diagnósticos são identificados quando a mesma sequência é encontrada em outros textos.

Esta mesma técnica foi adaptada para as necessidades do trabalho de Oliveira (2014) que identifica a continuidade de tratamento em resumos de alta. As fases e processos por ele utilizados são apresentados no Quadro 2. Com esta técnica ele obteve uma precisão de 93% e uma sensibilidade de 88% e *f-measure* de 90% utilizando as fórmulas:

$$precisão = \frac{VP}{(VP+FP)} \quad (1)$$

$$sensibilidade = \frac{VP}{(VP+FN)} \quad (2)$$

$$f\text{-measure} = 2 * \frac{precisão * sensibilidade}{(precisão + sensibilidade)} \quad (3)$$

Onde VP são os verdadeiros positivos, FP são os falsos positivos e FN são os falsos negativos, respectivamente, as identificações corretas, identificações incorretas e as continuidades de tratamento que não foram identificadas.

As técnicas citadas são muitas vezes utilizadas em conjunto com algum tipo de dicionário. Friedman (2000) mostra como foi a evolução do sistema MedLEE (Medical Language Extraction and Encoding), que utiliza tanto regras gramaticais como um *lexicon* (que pode ser interpretado como um dicionário, pois é um vocabulário da área médica). Neste trabalho é mencionado que para expandir o *lexicon* da área original de radiologia de tórax para a área de radiologia em geral levou uma pessoa ano, e para outras áreas subsequentes utilizou o esforço de uma pessoa por cerca de três meses. Vale salientar que no sistema MedLEE, tanto as regras gramaticais utilizadas como o *lexicon* são para a língua inglesa.

Dos trabalhos citados neste item, os de Bui *et al.* (2010), Bulegon (2011) e Oliveira (2014) têm maior similaridade técnica com o presente trabalho, pois utilizam regras de classificação gramatical, especialmente o último, que serviu de base para o processo utilizado, tendo poucas modificações.

Quadro 2 – Fases e processos utilizados por Oliveira (2014)

Fase	Processo	Descrição
Pré-processamento	Expansão de acrônimos	Os acrônimos são expandidos conforme uma tabela pré-definida, por exemplo IAM é expandido para infarto agudo do miocárdio.
	Remoção de caracteres especiais	Foram removidos os caracteres [-!?!><= %;/#,@*]
	Normalização para letras minúsculas	Todo o texto é transformado em letras minúsculas.
	Remoção de <i>stop-words</i> <sup>1</sup>	Os artigos são removidos.
	Formatação das datas	As datas são escritas por extenso.
PLN	Detecção de sentenças	Segmentação do texto em sentenças.
	Tokenização <sup>2</sup>	Separação dos termos e pontuações em uma única unidade.
	Detecção de nomes próprios	Identifica os nomes próprios compostos em uma única unidade.
	Expansão de preposições	Aglutinações como “da” são transformadas em “de a”.
	Etiquetagem morfológica	Cada unidade identificada no processo de tokenização recebe uma etiqueta morfológica.
Busca dos padrões das regras no texto	Inversão das sentenças.	A ordem das sentenças é invertida.
	Remoção das sentenças posteriores a quarta	Como a ordem está invertida, são eliminadas da primeira até a quarta sentença.
	Identificação das regras nos textos	Cada sequência de classificação morfológica que corresponde a uma regra é procurada no texto, e quando encontrada é considerada uma identificação. Este processo também é chamado de execução das regras.

<sup>1</sup> Palavras muito comuns e sem importância para o processo.

<sup>2</sup> Um Token é um segmento de texto ou símbolo que pode ser manipulado por um processo.

Textos médicos são fonte de informações para os pesquisadores na área de PLN, sendo que o laudo radiológico contém informações extremamente relevantes (Kahn, Langlotz e Burnside, 2009). Assim sendo os laudos radiológicos são um dos textos de interesse nas pesquisas que utilizam PLN.

## 2.2 PLN EM LAUDOS RADIOLÓGICOS

O laudo radiológico é um documento cujo objetivo primário é comunicar o resultado de um procedimento de imagem para o médico que o solicitou. Ele deve mencionar a indicação clínica, as técnicas e procedimentos utilizados e as possíveis intercorrências ocorridas durante o procedimento. A descrição dos achados deve ser redigida de forma sucinta e completa, e quando o diagnóstico é seguro, terminar com uma impressão/conclusão (Fenelon, 2003; Kahn, Langlotz e Burnside, 2009).

Os laudos radiológicos têm sido elaborados em texto livre desde o surgimento da radiologia, pois era o único meio disponível na época, e este aspecto não evoluiu (Reiner, Knight e Siegel, 2007).

O laudo radiológico pode ser um meio de troca de experiências entre diferentes médicos ou ser utilizado em pesquisas retrospectivas para preparação de material com o objetivo educacional ou de treinamento. Também permite a recuperação de determinado tipo de achado por médicos residentes (Kahn, Langlotz e Burnside, 2009).

Sociedades de especialidades médicas defendem a utilização de textos estruturados nos laudos. Estes são criados com formatos de texto e termos pré-definidos, nos quais os tópicos de um texto iniciam com termos padronizados para facilitar sua localização, e a ordem dos mesmos também é padronizada (Kahn, Langlotz e Burnside, 2009).

A recuperação da informação, citada anteriormente, é facilitada quando os textos são estruturados, e viabiliza a recuperação dos textos para pesquisa, ensino e para a melhora da qualidade clínica. Os laudos estruturados também facilitam a prática do radiologista, cobrança, controle de qualidade, concordância com as

regulamentações e documentação do serviço radiológico (Kahn, Langlotz e Burnside, 2009).

Recentemente têm sido realizadas pesquisas na recuperação de informação em laudos radiológicos. Foram selecionados as 10 mais recentes que contemplam uma variedade de exemplos de informações que hoje em dia estão sendo recuperadas de laudos radiológicos, tanto em idioma estrangeiro como em português.

Borkar, Nikose e Chaturvedi (2013) desenvolveram um sistema de mineração de textos, onde uma das partes do processo consiste em utilizar técnicas estatísticas para estruturar os laudos identificando os AR, de forma que os outros módulos do sistema interagem com as imagens correspondentes.

Yetisgen-Yildiz *et al.* (2013) ressaltam a importância do acompanhamento das recomendações dos radiologistas especialmente em situações não agudas. Eles propõe um processo de recuperação de informação que identifica automaticamente as sentenças de recomendações dos radiologistas em laudos de imagens. De forma semelhante, Dutta *et al.* (2013) identificam recomendações de exames adicionais de imagem para acompanhamento.

Dublin *et al.* (2013) complementam a validação manual para identificação de pneumonia em radiografias de tórax. O processo por eles desenvolvido pretende otimizar a acurácia ou minimizar a revisão manual.

O trabalho de Do *et al.* (2012) mostra a possibilidade de se tirar vantagem dos textos gerados em tempo real, através de reconhecimento de voz na confecção de laudos referentes às fraturas em ossos. A partir destes textos se recupera classificações relevantes ou critérios de diagnósticos ou probabilidade de doenças, facilitando assim o trabalho do radiologista.

Também Percha *et al.* (2012) recuperam informações bem específicas nos laudos de mamografias, onde os laudos são classificados pela categorização da composição dos tecidos dos seios.

Gerstmair *et al.* (2012) desenvolvem um sistema de pesquisa para recuperação de laudos e imagens baseado no texto dos laudos, construída para o idioma alemão. Este sistema realiza a indexação dos laudos e utiliza internamente o Morphosaurus (Schulz *et al.*, 2002) com o objetivo de permitir a recuperação de

termos semanticamente semelhantes. Neste trabalho ele afirma que não existe sistema que extraia referências de uma imagem fazendo a associação ao seu contexto.

Lacson *et al.* (2012) mostram um conjunto de ferramentas que podem ser utilizados em pesquisas sobre textos de laudos, e que podem ser utilizados sem necessidade de customização ou programação. Achados específicos podem ser procurados e recuperados em laudos que contenham o mesmo termo ou termo semelhante, incluindo variações léxicas ou semânticas.

Em relação a pesquisas em português, (Serapião *et al.*, 2013) desenvolveram um sistema para a criação de um índice extraído de laudos radiológicos com o objetivo de obter uma alternativa local aos padrões já existentes em outros idiomas.

Honorato *et al.* (2008), que também trabalha com textos em português, criam uma estruturação formada por atributo/valor, sendo que o todas as frases utilizadas são atributos, e o valor, que pode ser 0 ou 1, indica a presença ou ausência de uma determinada frase em um determinado laudo.

Destes artigos, somente Lacson *et al.* (2012) têm preocupação com a usabilidade.

O Quadro 3 são apresentados os objetivos de PLN em cada um dos 10 estudos, a técnica utilizada para recuperação de informação e a quantidade de laudos analisados.

Dos trabalhos citados neste item, os de Borkar, Nikose e Chaturvedi (2013), Gerstmair *et al.* (2012), (Serapião *et al.*, 2013) e Honorato *et al.* (2008) têm similaridade nos objetivos com o presente trabalho, pois não recuperam informações específicas.

Quadro 3 - Técnicas e objetivos dos artigos com a quantidade de laudos utilizados

Artigo	Técnicas	Objetivos	Quant. de Laudos
(Borkar, Nikose e Chaturvedi, 2013)	Aprendizado de máquina, estatística	Recuperar achados médicos de laudos radiológicos	-
(Yetisgen-Yildiz <i>et al.</i> , 2013)	Aprendizado de máquina. Estatística e regras.	Recuperar as recomendações dos radiologistas	800
(Dutta <i>et al.</i> , 2013)	Algoritmo desenvolvido pelos pesquisadores. Pesquisa por termos.	Recuperar as recomendações de exames adicionais	3264
(Do <i>et al.</i> , 2012)	Baseado em regras	Mostrar em tempo real classificações relevantes/critérios de diagnósticos/probabilidade de doenças	368
(Dublin <i>et al.</i> , 2013)	Utiliza o sistema ONYX <sup>1</sup> , baseado em regras gramaticais e redes bayseanas	complementar a validação manual para identificação de pneumonia em radiografias de tórax	5000
(Percha <i>et al.</i> , 2012)	Baseado em regras com <i>stem</i> <sup>2</sup>	Classificar laudos pela composição dos tecidos	600
(Gerstmair <i>et al.</i> , 2012)	Baseado em regras	Indexar dos laudos para futura recuperação	258.824
(Lacson <i>et al.</i> , 2012)	Baseado em dicionário e queries.	Recuperar achados específicos	338
(Serapião <i>et al.</i> , 2013)	Estatístico	Criação de um Índice para os laudos	61.461
(Honorato <i>et al.</i> , 2008)	Baseado em regras	Estruturação dos laudos pela identificação das frases que o compõe	-

<sup>1</sup> O ONYX é um sistema de análise semântica de textos clínicos.

<sup>2</sup> Stem é uma redução da palavra para sua raiz, mas sem o compromisso de representar a raiz morfológica da palavra.

### 3 MÉTODO

Trata-se de uma pesquisa exploratória e metodológica, pois a técnica utilizada na recuperação de informação nos laudos não estava completamente determinada no início dos trabalhos e foi proposta não somente a técnica, como também um método para utilizá-la. A abordagem é quantitativa tendo em vista que a adequação ou não da técnica e do método propostos foi identificada com base nos valores das proporções de acerto e erro na utilização dos mesmos.

Quadro 4 – Fases e etapas do trabalho

	Fases		Etapas
1	Desenvolvimento do protótipo do sistema ESTEMED	1.1	Reunião inicial com a equipe de desenvolvimento
		1.2	Organização do material e reunião com a equipe
		1.3	Definição das funcionalidades e criação dos diagramas
		1.4	Versões do protótipo do ESTEMED e projeto voltado ao usuário
		1.5	Utilização do protótipo pela equipe de desenvolvimento.
2	Criação e execução das regras	2.1	Criação das regras
		2.2	Execução das regras
3	Utilização do ESTEMED pela equipe de teste		
4	Proposição do método		

Para identificação das classes morfológicas das palavras foram escolhidas as bibliotecas da ferramenta de PLN CoGrOO, pois é bem adaptada ao português. Foi utilizada a linguagem de programação Java no desenvolvimento do protótipo, uma vez que as referidas bibliotecas também estão nesta linguagem.

A metodologia de desenvolvimento do protótipo do ESTEMED seguiu os passos tradicionais do ciclo de vida de desenvolvimento de sistemas utilizando modelos da *Unified Modeling Language* (UML), notadamente os diagramas de

classes e de sequência. Foi utilizado o conceito de prototipação evolutiva, que consiste em desenvolver um protótipo do sistema e iniciar sua utilização na qual se realimenta o processo de desenvolvimento com informações para que o protótipo possa evoluir adequadamente.

Esta pesquisa foi elaborada em quatro fases apresentadas com suas etapas no Quadro 4.

### 3.1 BASE DE DADOS

Quanto a amostra de laudos utilizados neste processo, eles foram extraídos das bases do Sistema de Informação em Radiologia do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo. A utilização destes dados foi aprovada por meio do processo número 4199/2010 do comitê de ética em pesquisa do Hospital das Clínicas de Ribeirão Preto. Foram selecionados os laudos de Ressonância Magnética de Encéfalo (período 21/01/1999 e 05/02/2009) totalizando aproximadamente 23.000 laudos. No Quadro 5 são apresentadas as quantidades utilizadas para cada atividade.

Quadro 5 – Quantidades de laudos utilizados em cada atividade.

Testes Iniciais	Base de teste da equipe de desenvolvimento	Base de treinamento da equipe de desenvolvimento	Base de teste da equipe de teste	Base de treinamento da equipe de teste
50	50	200	50	50

Para a criação dos laudos no sistema, o radiologista pode utilizar uma lista de frases comuns, que indicam situações de normalidade. Quando é o caso, o radiologista simplesmente utiliza a frase, e quando é necessário, a complementa com as alterações necessárias. Observa-se no conjunto de textos, no entanto, que a grande variedade de situações individuais obriga o radiologista a alterar aquelas frases utilizadas como base de forma tão intensa, que as mesmas deixam de ter utilidade sob o ponto de vista de estruturação. Também vale lembrar que as frases comuns utilizadas não estão no formato de estruturação proposto no presente trabalho.

Para garantir uma melhor qualidade nos textos, os mesmos são criados por profissionais em início de carreira e revisados por médicos radiologistas experientes.

Os laudos foram manipulados em planilha eletrônica de forma a serem randomizados pela função “ALEATÓRIO” do editor de planilha eletrônica “LibreOffice calc” para evitar influência da ordem original dos mesmos na pesquisa.

As bases de treinamento e teste foram extraídas desta planilha já randomizada à medida da necessidade do desenvolvimento e dos testes.

### 3.2 LOCAL DO ESTUDO

O desenvolvimento foi realizado no Laboratório de Informática em Saúde do Programa de Pós-Graduação em Tecnologia em Saúde da Escola Politécnica da Pontifícia Universidade Católica do Paraná.

### 3.3 PARTICIPANTES

São dois grupos que participaram da pesquisa, um no desenvolvimento e outro na avaliação. A equipe de desenvolvimento foi formada por dois acadêmicos do curso de engenharia de computação com experiência em Java e na utilização das bibliotecas do CoGrOO, o pesquisador, que é analista de sistemas experiente, que juntos realizaram a modelagem e codificação do ESTEMED e uma acadêmica de medicina que utilizou as primeiras versões do protótipo e colaborou na evolução do mesmo e da metodologia para a sua utilização. Os acadêmicos participaram de projetos de iniciação científica e projetos de iniciação em desenvolvimento tecnológica neste grupo de pesquisa.

A equipe de testes foi formada por quatro médicos neurologistas e três enfermeiras, todos do Programa de Pós-Graduação em Tecnologia em Saúde da Escola Politécnica da Pontifícia Universidade Católica do Paraná participantes do grupo de pesquisa de processamento de imagens médicas.

### 3.4 O DESENVOLVIMENTO DO PROTÓTIPO DE SISTEMA ESTEMED

Realizado na fase 1 deste trabalho, neste desenvolvimento foram adotadas 5 etapas (Quadro 4), contando com a participação da equipe de desenvolvimento e de teste.

Este trabalho teve como base o algoritmo de Oliveira (2014), pois o mesmo se mostrou adequado à recuperação de informação utilizando uma estratégia cuja adaptação a uma ferramenta de fácil utilização por profissionais de saúde se mostrou plausível e foi uma das recomendações do autor para trabalhos futuros.

Na primeira etapa, reunião inicial, foram levantadas as funcionalidades. A estratégia desta reunião foi apresentar as funcionalidades do algoritmo de Oliveira (2014) e identificar quais funcionalidades gráficas seriam necessárias para possibilitar uma fácil manipulação dos dados, criação e execução das regras. Foram considerados aspectos de usabilidade, de forma que profissionais de saúde pudessem facilmente entender e executar as funcionalidades propostas.

Para a identificação das funções gráficas que seriam criadas, foram enfatizados dois processos principais, a criação e a execução das regras. O processo de criação de regras é caracterizado por alguns passos principais. Primeiro o texto deve ser selecionado, como por exemplo o texto da Figura 1:

Figura 1 – texto de exemplo para mostrar algoritmo

Assimetria de dimensões do lobo temporal menor à esquerda com áreas de cavitação na região amigdaliana com borramento da substância branca nessa região. Hipocampos de dimensões assimétricas, menor à esquerda, observando-se hipersinal no hipocampo e giro parahipocampal deste lado. Acentuação dos sulcos cerebrais na região parietal alta.

Ventrículos cerebrais, cisternas e fissuras dentro da normalidade. Acentuação de sulcos cerebelares. Tronco cerebral sem alterações. Ausência de coleções extra axiais.

A seguir, deve-se escolher um trecho que é de interesse, como por exemplo “Assimetria de dimensões” que é parte de um AR. Deve-se então associar este

trecho a uma regra, que terá a sequência de etiquetas “[N\_F\_S\_] [PRP\_] [N\_F\_P\_]” que representam estas palavras, pois a primeira palavra é um substantivo feminino plural, a segunda palavra é uma preposição e a terceira palavra é um substantivo feminino plural.

Com a execução desta regra temos como exemplos de identificações “presença de saliências”, “expansão de lesões” e “irregularidade de placas” todos partes de AR com as mesmas sequências de etiquetas. Estes resultados devem ser apresentados de forma intuitiva para os usuários, de forma que facilite a identificação e marcação dos FP e FN.

O material desta reunião foi organizado com a criação de uma tabela que teve as funcionalidades do protótipo em uma coluna e a descrição das mesmas em uma segunda coluna. Foi realizada outra reunião com equipe de desenvolvimento e de teste, na qual foi apresentado o resultado da primeira e foram debatidas todas as funcionalidades que seriam incorporadas. As funcionalidades foram mantidas e aprovadas no projeto baseado no consenso entre os participantes.

Após esta reunião a equipe de desenvolvimento formalizou todas as funcionalidades aprovadas e desenvolveu os diagramas de classes e objetos e os diagramas de sequência iniciais.

Foi criado então um protótipo do ESTEMED baseado no processamento de PLN utilizado por Oliveira (2014) conforme descrito no Quadro 2 do capítulo 2, com a eliminação de dois processos da fase 3, porque só se aplicavam a situação específica daquela pesquisa: a inversão de sentenças e a remoção de sentenças posteriores a quarta. A parte gráfica foi desenvolvida a partir dos diagramas que foram criados, utilizando a linguagem de programação Java e o ambiente de desenvolvimento NETBEANS. O banco de dados utilizado foi o MySQL, pois não tem custo de licença, já havia experiência na sua utilização pela equipe e tecnicamente não havia restrições.

A partir deste ponto foi feita a prototipação evolutiva, baseada na utilização do sistema para realização dos testes iniciais, realizados pela própria equipe de desenvolvimento, sempre com a preocupação de facilitar a utilização por profissionais de saúde.

A primeira versão do ESTEMED utilizou diretamente a planilha onde estavam os textos, mas logo foi considerado mais conveniente manipular os textos em banco de dados e assim foi implementada a funcionalidade de importação de arquivos para o MySQL, onde já estavam os usuários e as regras.

Também foi levantada a possibilidade de uma identificação de sequência de etiquetas morfológicas serem confirmadas por outra sequência, e assim foi proposta e implementada a funcionalidade de sub-regras. Com esta facilidade tornou-se possível criar uma sub-regra associada a uma das regras, de forma que a segunda só seria considerada caso a primeira fosse encontrada.

Nos testes iniciais foram identificadas algumas situações nas quais o processamento que estava sendo utilizando mostrou problemas, sendo então necessário retirar três processos da técnica utilizada por Oliveira (2014): a remoção de caracteres especiais, a detecção dos nomes próprios e a remoção das *stop-words*.

Em alguns casos foi notado que a ausência das vírgulas dificultava a identificação dos textos correspondentes, e a regra sem a vírgula passava a gerar falso FP. Com a retirada do processo que elimina caracteres especiais estes FP deixaram de ocorrer.

No caso dos nomes próprios, foi observado que nenhum dos laudos utilizados tinha nomes próprios, então este processo seria inútil. Com este processo, as diversas estruturas anatômicas citadas nos laudos que são nomes de pessoas não estavam sendo identificadas.

Nas tentativas de melhorias das regras, havia a possibilidade de se discernir melhor entre AR e TM através da presença ou ausência de artigos. Desta forma o processo de remoção de *stop-words*, também foi eliminado.

Além dos processos retirados também foi necessário criar um processo a mais que identificasse se a frase somente continha observações de normalidade. Ou seja, frases que não identificavam nenhuma anormalidade. Estas frases foram identificadas pela presença de trechos, como “sem sinal”, “dentro da normalidade”, “não apresenta” e “aspecto normal” e portanto não foram consideradas no resto do processo.

Nos testes iniciais pela equipe de desenvolvimento, as atividades seguiram uma ordem mais intuitiva, seguindo as prioridades das atividades de desenvolvimento do ESTEMED. Não houve preocupação, naquele momento, com a ordem e a sistematização das atividades, sendo apenas anotado o que foi feito para posterior criação de um método.

### 3.5 CRIAÇÃO E EXECUÇÃO DAS REGRAS

A criação e execução das regras correspondem a fase 2 do trabalho. Ao longo deste processo foram observadas situações em que a realização desta tarefa era dificultada. A partir disso, foram realizadas alterações no processo, objetivando elaborar um método que facilite a criação de regras. O desenvolvimento desse método corresponde à fase 4 do trabalho.

Os resultados da criação e execução das regras são apresentados no item 4.2 “Resultados da Aplicação do ESTEMED” do capítulo 4.

#### Criação das regras

Com base na experiência inicial obtida com a utilização do ESTEMED em seus testes iniciais e nas anotações das atividades realizadas durante o desenvolvimento do sistema, utilizamos os seguintes passos:

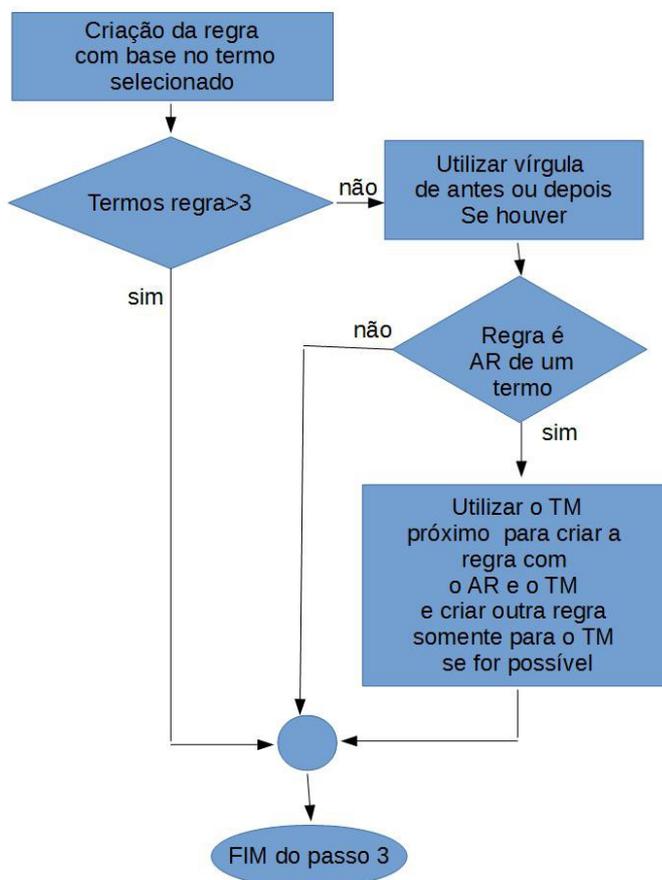
**Passo 1:** Randomização da base de laudos e criação das bases de teste e de treinamento. Com o objetivo de evitar qualquer viés associado à ordem dos laudos na base foi realizada a randomização utilizando a função “ALEATÓRIO” do editor de planilhas do LibreOffice. Já haviam sido utilizados em testes iniciais 50 laudos, então foram selecionados os próximos 250, separando os primeiros 50 para base de treinamento e os outros 200 para base de teste. Estas quantidades foram escolhidas tendo em vista o esforço que foi necessário durante os testes iniciais e que um dos aspectos que orientaram a criação do método foi evitar uma utilização mais significativa de recursos humanos.

**Passo 2:** Como os participantes da equipe de desenvolvimento que realizaram estes testes não têm formação na área de saúde, foi solicitado a 5

participantes da equipe de teste, quatro médicos e uma enfermeira, a marcação dos TM e AR, tendo em vista o seu amplo conhecimento no assunto. Cada um deles marcou manualmente 100 laudos que foram impressos com este fim, de forma que cada grupo de 50 laudos foi marcado por duas pessoas, com o objetivo de se poder fazer uma comparação em caso de dúvida.

**Passo 3:** Com base na marcação feita em papel, a equipe de desenvolvimento, utilizando o ESTEMED, marcou os termos na base de treinamento, sempre em conjunto com pelo menos mais uma palavra. Esta marcação procurou identificar padrões que indiquem os TM ou AR marcados em papel pela equipe de teste, seguindo a mesma lógica adotada no método de Oliveira (2014). Com esta marcação, identificando se era um TM ou AR, as regras foram criadas. Nos testes iniciais, foi percebido que algumas orientações podiam ser utilizadas na criação das regras. Assim, a heurística utilizada foi de sempre que possível utilizar pelo menos três palavras, e quando não foi possível, tentar utilizar vírgula antes ou depois. Não houve restrição em relação ao número máximo de palavras. Os AR que continham somente uma palavra foram associados com o TM próximo e a regra foi criada como AR. Outra regra específica para o TM usado foi criada quando possível. Estas orientações foram sendo adotadas ao longo dos testes iniciais pela observação das regras que tinham resultados melhores ou piores, ou seja, procurando maximizar a quantidade de VP e minimizar a quantidade de FN. Na Figura 2 é apresentado o fluxograma deste passo. No Quadro 6 temos um trecho extraído de um laudo radiológico, a regra obtida do mesmo, e outro texto encontrado com base nesta regra. No Quadro 7 temos todos as etiquetas utilizadas nas regras e seus significados.

Figura 2 – Fluxograma do passo 3



Quadro 6 - Utilização de regra em laudo radiológico

Trecho original	dilatação dos ventrículos
Regra gerada	[N_F_S_][PRP_][DET_M_P_][N_M_P_]
Trecho encontrado em outro texto	ectasia de os espaços

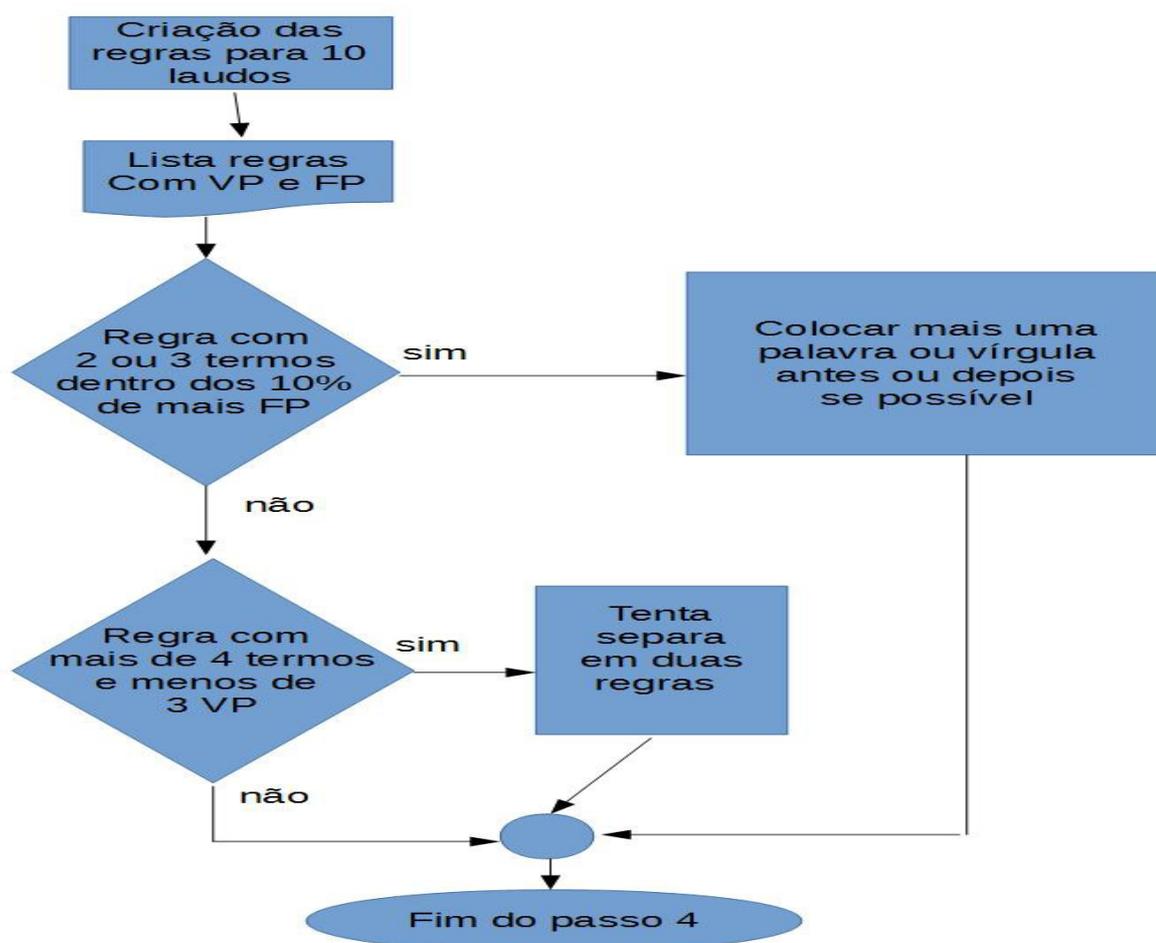
Quadro 7 - Etiquetas utilizadas e seus significados

-PNT_NSEP	,
-PNT_REL	( ou )
ADJ_F_P_	Adjetivo Feminino Plural
ADJ_F_S_	Adjetivo Feminino Singular
ADJ_M/F_S_	Adjetivo Singular
ADJ_M_P_	Adjetivo Masculino Plural
ADJ_M_S_	Adjetivo Masculino Singular
ADV_	Advérbio
DET_F_P_	Artigo Definido Feminino Plural
DET_F_S_	Artigo Definido Feminino Singular
DET_M_P_	Artigo Definido Masculino Plural
DET_M_S_	Artigo Definido Masculino Singular
KC_	Conjunção Coordenada
KS_	Conjunção Subordinativa
N_F_P_	Substantivo Feminino Plural
N_F_S_	Substantivo Feminino Singular
N_M_P_	Substantivo Masculino Plural
N_M_S_	Substantivo Masculino Singular
PRP_	Preposição
V_GER_	Verbo no Gerúndio
V_PCP_	Verbo Particípio
V_PCP_F_P_	Verbo Particípio Feminino Plural
V_PCP_F_S_	Verbo Particípio Feminino Singular
V_PCP_M_P_	Verbo Particípio Masculino Plural
V_PCP_M_S_	Verbo Particípio Masculino Singular

**Passo 4:** Após a criação de regras para cada conjunto de 10 textos, foi realizada uma execução dessas regras no ESTEMED nos 50 laudos da base de treinamento para verificar o resultado gerado com elas, identificando os VP e FP. A partir disto foi criada uma lista com a quantidade de VP e FP de cada regra. As que obtiveram maior quantidade de FP ou maior quantidade de FP do que VP foram analisadas para identificar Nas regras com dois ou três termos que tinham as maiores quantidades de FP (aproximadamente os

10% que em possibilidades de melhoramentos. Nas regras com mais de 4 termos que tinham menos de 3 VP era feita uma tentativa de separação dos termos para obter-se duas regras com menos termos. cabeçavam a lista) era procurada uma palavra ou pontuação para aumentar a quantidade de termos na regra e torná-la mais específica. Na Figura 3 é apresentado o fluxograma deste passo.

Figura 3 – Fluxograma do passo 4



Esta execução e análise dos resultados foi realizada de 10 em 10 textos por três motivos. Primeiro, observou-se nos testes iniciais que usuários com pouca experiência no assunto tendem a criar regras que geram um grande número de FP ou que têm poucos VP e com esta avaliação simples a curva de aprendizado é acelerada. Segundo, o usuário tende a perder a concentração quando chega ao final de um número muito grande de textos, e com esta atividade, a concentração não é perdida, pois as regras de 10 textos são criadas em aproximadamente uma hora. Terceiro, se a atividade de criação de regras ocorrer num período muito menor que uma hora, a produtividade tende a ser menor devido as interrupções constantes. <sup>1</sup>

**Passo 5:** Com todas as regras criadas para os 50 laudos de treinamento, foram então marcados todos os FP e os FN utilizando a tela de resultados do ESTEMED. Foram considerados como VP quando uma regra encontrou um trecho que corresponde ao mesmo elemento da regra (AR ou TM) ou que tem parte dele que corresponda. Caso contrário foi considerado FP. Qualquer palavra ou conjunto de palavras que faça parte de um AR ou TM e que não foi encontrado por nenhuma regra foi considerado FN. Com base na lista de todos os FN foi realizada mais uma tentativa de criação de regra para cada um deles, e para as regras com mais FP do que VP foi feita mais uma tentativa de ajuste, verificando qual a palavra anterior e posterior para todas as ocorrências, procurando por alguma que discernisse entre os VP e os FP.

**Passo 6:** Procurar nos FP a etiqueta morfológica da palavra anterior ou posterior que os caracterize. Só foi encontrado o caso de artigos antes dos FP de uma regra de AR. Desta forma estes FP puderam ser eliminados.

**Passo 7:** Para todos os FP e FN que a classificação morfológica da palavra foi errada (feita pelo CoGrOO, devido a limitação do *corpus* anotado) foi criado registro em uma tabela do banco de dados com a palavra e a classificação correta. O ESTEMED foi alterado para que quando a palavra estivesse nesta tabela fosse utilizada a classificação da tabela e não a do CoGrOO. Foram colocadas 120 palavras nesta tabela.

---

<sup>1</sup>Nos passos 4, 6, 7 e 8 é utilizando acesso direto à base de dados, pois o ESTEMED não possui ainda estas funcionalidades.

**Passo 8:** Quando dois resultados são de mesmo elemento (AR ou TM) e um está completamente contido no outro, o primeiro deve ser eliminado. Isso é feito para evitar distorções nos cálculos.

### **3.5.1 Execução das regras**

Com as regras criadas e sem encontrar-se possibilidades de melhorias, as mesmas foram aplicadas na base de teste. Os passos 6,7 e 8 também foram aplicados à base de teste. O resultado foi analisado e as quantidades encontradas corretamente (VP), as não encontradas mas que deveriam ter sido encontradas (FN) e a encontradas que não deveriam ter sido encontradas (FP) foram marcadas e anotadas. Os cálculos dos valores de precisão e sensibilidade seguiram as mesmas fórmulas que Oliveira (2014).

Os resultados estão compilados no item 4.2.1 “Utilização pela equipe de desenvolvimento”.

## **3.6 UTILIZAÇÃO DO ESTEMED PELA EQUIPE DE TESTE**

Para a fase 3, duas enfermeiras e três médicos neurologistas da equipe de testes criaram as regras utilizando 10 laudos cada. Três enfermeiras dividiram 50 laudos na identificação dos FP e FN nos resultados. Antes de realizar a criação das regras a equipe foi treinada por 30 minutos. Neste tempo foi mostrado o ESTEMED, as atividades que seriam desenvolvidas e os passos do método que seriam aplicados. A maior parte do tempo despendido no treinamento não foi com as funcionalidades do ESTEMED, mas sim com relação às recomendações de quando e como criar as regras.

Entre a criação das regras e a identificação de FP e FN a equipe de desenvolvimento realizou a consolidação das regras em um usuário só, pois cada um tinha seu cadastro de usuário no ESTEMED e o sistema não permite que um usuário utilize as regras de outro. Também depois da criação das regras a equipe de desenvolvimento realizou a execução das mesmas. Foi encontrado nos resultados

15 casos de palavras classificadas erroneamente pelo CoGrOO. Estas palavras e suas classificações corretas foram colocadas na tabela correspondente. Foi realizada uma outra execução das regras gerando o resultado utilizado na identificação de FP e FN.

Após a identificação dos FP e FN a equipe de desenvolvimento realizou uma revisão na identificação dos FN, marcando mais 82 casos como FN.

É importante enfatizar que esta fase tem como objetivo a verificação na facilidade do uso do ESTEMED pelos profissionais de saúde, o método aplicado foi simplificado ao máximo, e o primeiro resultado obtido não passou por análise das regras com o objetivo de aumentar os VP e diminuir os FN e FP.

### 3.7 PROPOSIÇÃO DO MÉTODO

Este item corresponde a fase 4 do trabalho e têm como resultado o método apresentado no item 4.3 “MÉTODO PROPOSTO”. Para propor o método de estruturação dos laudos de ressonância magnética de encéfalo foi utilizada a experiência adquirida ao longo da utilização do ESTEMED, com o objetivo de estruturar os laudos radiológicos através da identificação e diferenciação nos trechos dos laudos classificados como TM e AR, obtendo assim a estrutura “o que foi encontrado” e “onde foi encontrado”

Com os ajustes realizados no protótipo do sistema ESTEMED, tendo em vista as características dos textos e dos resultados obtidos, e com base no método utilizado no item 3.5.1 “Criação das regras”, foi proposto um método para criação de regras para possibilitar a recuperação das informações pretendidas pelo usuário nos textos por ele utilizados, que está apresentado no item 4.3 “MÉTODO PROPOSTO”. Este método é basicamente o utilizado pela equipe de desenvolvimento, com pequenas generalizações, adaptações e simplificações. Assim, situações muito específicas, como qual a forma de randomização, foram eliminadas. Também passos que estavam complexos foram divididos, gerando passos a mais, como na criação de regras para os FN que ficou em um único passo.

## 4 RESULTADOS

A apresentação dos resultados aborda o protótipo ESTEMED, o método proposto para sua utilização e os resultados da utilização do ESTEMED pela equipe de desenvolvimento e pela equipe de teste.

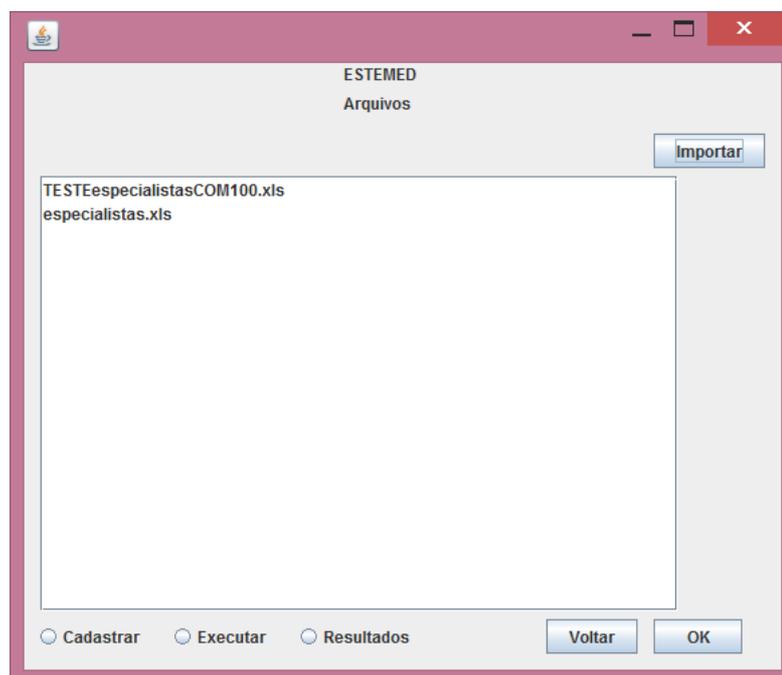
### 4.1 O ESTEMED

O protótipo desenvolvido possui funcionalidades de criação e identificação de usuários, importação de textos, criação de regras, execução e apresentação de resultados das execuções das regras.

Na tela de entrada do sistema é possível cadastrar um novo usuário ou utilizar o login de um já criado. Após o login é apresentada a lista de arquivos já importados. Através do botão “Importar” pode-se importar o arquivo desejado, que passa então a fazer parte da lista de arquivos. A tela com a lista de arquivos é apresentada na Figura 4, onde pode-se selecionar o arquivo desejado e uma das operações disponíveis (“Cadastrar”, “Executar” e “Resultados”), e clicando no botão “OK” passa-se para a operação escolhida.

Se for escolhido “Cadastrar”, é mostrada a tela de cadastramento, mostrada na Figura 5, com a ordem das ações necessárias para se cadastrar uma regra, que são: selecionar o texto na lista, selecionar o trecho do texto, selecionar o tipo de texto, TM ou AR, clicar em “criar regra”. A função de criação de sub-regras está disponível, conforme foi colocado no item 3.4 “O Desenvolvimento do protótipo de sistema ESTEMED”. No entanto, o item 3.5.1 “Criação de Regras” esta facilidade não foi utilizada, pois não foram encontradas situações nas quais fosse útil. As regras criadas são associadas a um “elemento”, que corresponde a um tipo de texto, que pode ser AR ou TM, e a um conjunto padrão de teste, que é uma forma de agrupar as regras de forma arbitrária.

Figura 4 – Tela com a lista de arquivos.



Se for escolhido “Executar” será mostrada a tela da Figura 6 com as regras existentes, para que as mesmas sejam selecionadas para esta execução. Esta seleção pode ser feita regra por regra na coluna “Seleção”, ou utilizando uma seleção prévia somente com as regras de um conjunto, de um elemento ou de uma execução anterior, através das listas na parte superior da tela. O “Conjunto” de uma regra pode ser alterado com o botão “Trocar Conjuntos”. Cada execução pode ser associada a uma descrição se o campo “Descrição da Execução” for preenchido. Este campo foi criado para facilitar a identificação posterior de uma execução. Também é possível passar para a tela de execução clicando no botão “Executar” da tela de cadastramento de regras, mostrada na Figura 5.

Após se clicar OK na tela da Figura 6 aparecerão os resultados, e podemos ter na sequência as seguintes ações, destacadas na Figura 7: escolher o texto, selecionar o trecho que corresponde a um AR ou TM não identificado por nenhuma regra e clicar no botão “Marcar Falso Negativo”; colocar na coluna observação quando um achado é FP.

Figura 5: Tela de cadastramento de regras

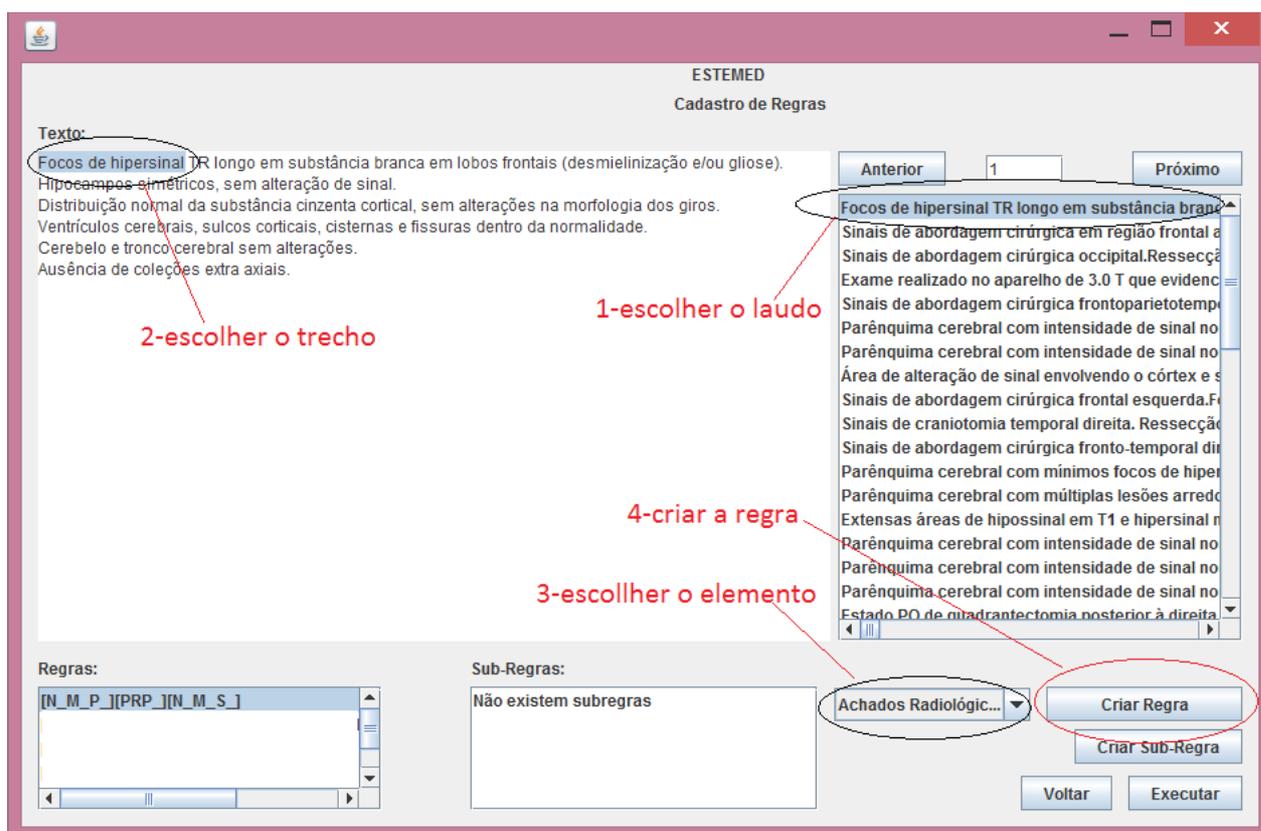


Figura 6 – Tela de escolha das regras para execução

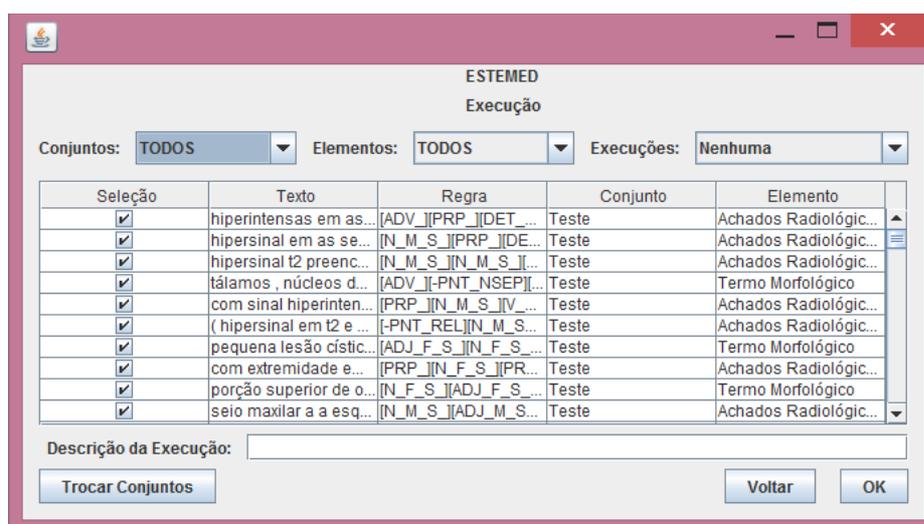


Figura 7 – Tela de resultados

The screenshot shows the ESTEMED software interface. At the top, there are buttons for "Marcar Falso Negativo" and "Cadastrar Regra". The main text area contains a medical report snippet with several words highlighted in different colors: blue (morphological terms), green (radiological terms), and red (morphological and radiological terms). Annotations in red text point to these highlights: "3-criar falsos negativos" points to the "Marcar Falso Negativo" button; "1-escolher o laudo" points to the "Voltar" button; "2-marcas os falsos negativos" points to the highlighted text; "azul=termo morfológico", "vermelho=achado radiológico e termo morfológico", and "verde=achado radiológico" explain the color coding. A table at the bottom lists the identified elements with columns for idRegra, Regra, Trecho Encontrado, Nome do Elemento, No Frase, PosInicial, PosFinal, and Comentário. A red circle highlights the "70 FP" in the "Comentário" column of the row for "efeito expansivo sobre", with the annotation "4-marcas falsos positivos".

**3-criar falsos negativos**

**1-escolher o laudo**

**2-marcas os falsos negativos**

azul=termo morfológico

vermelho=achado radiológico e termo morfológico

verde=achado radiológico

**4-marcas falsos positivos**

idRegra	Regra	Trecho Encontrado	Nome do Elemento	No Frase	PosInicial	PosFinal	Comentário
1522	[PRP ][DET_M_S ][N_M_S ][A..	de o lobo temporal	Achados Radiológicos	1	59	62	FP
1516	[PRP ][DET_M_S ][N_M_S ][A..	de o lobo temporal	Termo Morfológico	1	59	62	
1633	[N_M_S ][ADJ_M_S ][PRP ]	efeito expansivo sobre	Termo Morfológico	2	68	70	FP
1585	[N_M_S ][ADJ_M_S ]	efeito expansivo	Achados Radiológicos	2	68	69	
1542	[PRP ][DET_M_P ][N_M_P ][A..	sobre os ventrículos laterais	Termo Morfológico	2	70	73	
1434	[N_M_P ][ADJ_M_P ]	ventrículos laterais	Achados Radiológicos	2	72	73	FP
1593	[ADJ_M_S ][N_M_S ]	terceiro ventrículo	Termo Morfológico	2	75	76	
2015	[N_M_S ][-PNT_NSEP]	ventrículo	Termo Morfológico	2	76	77	
1790	[DET_F_S ][N_F_S ][ADJ_F_S ]	a linha média	Termo Morfológico	2	79	81	
1438	[N_F_S ][ADJ_F_S ]	linha média	Achados Radiológicos	2	80	81	FP
1667	[N_M_S ][PRP ][N_M_S ]	reforço pós contraste	Termo Morfológico	2	89	91	FP
1703	[N_M_S ][PRP ][N_M_S ]	reforço pós contraste	Achados Radiológicos	2	89	91	

Os resultados também são apresentados se na tela da Figura 4 for selecionado “Resultados”. Os trechos que foram identificados por alguma regra estão em negrito e cores, de forma que quando um trecho corresponde a uma regra de TM fica em azul, quando corresponde a uma regra de AR fica em verde e quando corresponde as duas, fica em vermelho. Uma vez marcado um FN ele fica em destaque.

Na parte inferior da tela há uma lista com todas as identificações obtidas pelo algoritmo de PLN no texto da parte superior esquerda da tela. As colunas e seus significados são:

- “idRegra” → número único associado a regra
- “Regra” → sequência de etiquetas desta regra
- “Trecho Encontrado” → trecho identificado pela regra

- “Nome do Elemento” → tipo de informação desta regra e pode ser “AR” ou “TM”
- “No Frase” → número da frase onde o trecho foi identificado
- “PosInicial” → posição inicial do trecho encontrado
- “PosFinal” → posição final do trecho encontrado
- “Comentário” → campo que pode ser preenchido com qualquer comentário que o usuário desejar, e que foi utilizado para colocar a sigla “FP”.

Podemos observar que a etiqueta “-PNT-NSEP” e as iniciadas por “DET” não estão presentes no Quadro 1 pois a pesquisa de Bulegon (2011) eliminou a pontuação e os artigos, que respectivamente correspondem às etiquetas mencionadas.

## 4.2 RESULTADOS DA APLICAÇÃO DO ESTEMED

### 4.2.1 Utilização do sistema pela equipe de desenvolvimento

Na fase 2, etapa 1 foram criadas 182 regras utilizando o método descrito no item 3.5 “**CRIAÇÃO DAS REGRAS**”, na base de treinamento de 50 laudos previamente marcados em papel por especialistas. Destas, 45 não encontraram textos equivalentes na base de teste, ou seja, houve 45 trechos considerados de interesse na base de treinamento que geraram regras cujas sequências de classificações morfológicas não foram encontradas na base de teste. Estas não serão portanto consideradas nos resultados, a não ser na Tabela 1, que apresenta a soma da quantidade de regras criadas por grupo de 5 textos. O objetivo desta tabela é mostrar o aumento ou diminuição da quantidade de regras criadas ao longo do processo.

Na Tabela 2 são apresentados os resultados correspondentes às regras agrupadas pelo número de termos das mesmas. Por exemplo, a regra “[N\_M\_S\_] [N\_M\_S\_]” tem dois termos, a regra “[N\_F\_S\_] [PRP\_] [N\_M\_S\_]” tem três termos, e assim por diante. Com esta tabela pode-se observar qual a quantidade de termos

que geraram melhores resultados, em comparação com a quantidade de regras criadas com a respectiva quantidade de termos.

Tabela 1 – Equipe de desenvolvimento - Quantidade de regras por grupo de 5 textos.

textos		
De	à	Soma de regras
1	5	34
6	10	14
11	15	23
16	20	11
21	25	21
26	30	19
31	35	9
36	40	15
40	45	9
46	50	27

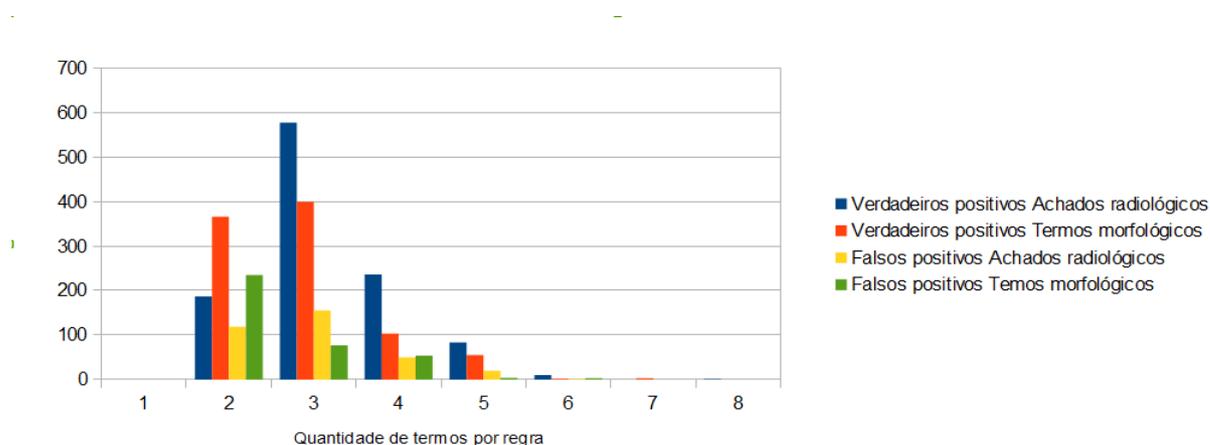
Tabela 2 - Equipe de desenvolvimento - Quantidade de regras por número de termos, com a quantidade de trechos identificados no texto e quantidade de VP e FP

Número de Termos por Regra	Total de Regras			VP			FP			Total		
	AR	TM	T	AR	TM	T	AR	TM	T	AR	TM	T
2	10	14	24	187	366	553	119	235	354	306	601	907
3	34	16	50	577	399	976	155	77	232	732	476	1208
4	20	16	36	236	103	339	50	54	104	286	157	443
5	9	6	15	83	55	138	20	4	24	103	59	162
6	7	1	8	10	2	10	1	3	4	11	3	14
7	0	2	2	0	3	3	0	0	0	0	3	3
8	1	0	1	1	0	1	0	0	0	1	0	1

A quantidade total de trechos identificados na base de teste foi 2737 sendo que houve 1332 casos de intersecções entre AR e TM, ou seja, partes de trechos identificados como AR e TM.

A Figura 8 mostra um gráfico que permite uma melhor comparação nos resultados obtidos pelo conjunto de regras, separadas pela quantidade de termos.

Figura 8 - equipe de desenvolvimento - VP e FP por quantidade de termos das regras



A Tabela 3 apresenta os valores de VP, FP, FN e os valores de sensibilidade e precisão. Na Tabela 4 é possível observar a variação na quantidade de identificações obtidas nos textos, pois ela apresenta a média e o desvio padrão dos VP e FP por texto, para AR, TM em conjunto nas primeiras colunas e na sequência os valores para AR e TM separadamente.

Tabela 3 – Equipe de desenvolvimento - Valores obtidos

<i>F-measure</i>	Precisão	Sensibilidade	VP	FP	FN
0,7713	0,7387	0,807	2016	713	482

Tabela 4 – Equipe de desenvolvimento - Média e desvio padrão de VP e FP por texto

	Média	Desvio Padrão	Média AR	Desvio Padrão AR	Média TM	Desvio Padrão TM
VP por Texto	12,14	7,95	6,69	4,78	5,60	3,88
FP por texto	4,16	3,14	2,42	1,59	2,54	1,85

Tabela 5 – Equipe de desenvolvimento - Resultados das melhores regras

Id	Composição	Treinamento		Teste	
		VP	FP	VP	FP
351	[N_F_S_][ADJ_F_S_][ADJ_F_S_]	14	0	96	10
363	[N_M_P_][ADJ_M_P_]	25	1	104	18
397	[N_F_P_][KC_][N_F_P_]	10	0	27	0
415	[N_F_P_][PRP_][N_F_S_]	12	0	12	2
435	[N_F_S_][PRP_][DET_M_P_][N_M_P_]	13	0	51	2
475	[N_M_S_][ADJ_M_S_][ADJ_M_S_]	12	2	62	8
483	[N_F_S_][PRP_][N_M_S_]	24	2	74	14

A Tabela 5 mostra o resultado apenas das regras que foram consideradas de performance superior na base de treinamento, objetivando verificar qual é a proporção das quantidades de identificações entre a base de treinamento e de teste. Esta seleção foi feita pegando as regras com maior número de VP e com até 3 FP. Também é mostrado o resultado destas regras na base de teste. A Tabela 6 mostra o resultado apenas das regras que foram consideradas de performance inferior na base de treinamento. Estas foram selecionadas quando obtiveram mais resultados FP do que VP. Também é mostrado o resultado destas regras na base de teste.

Tabela 6 – Equipe de desenvolvimento - Resultados das piores regras

Id	Composição	Treinamento		Teste	
		VP	FP	VP	FP
379	[N_M_P_][PRP_][N_M_S_]	1	2		
453	[N_F_S_][PRP_][N_F_P_]	1	2		
459	[N_F_P_][ADJ_F_P_]	4	6	7	33
563	[ADJ_M_S_][PRP_][DET_F_S_][N_F_S_]	3	6		
637	[ADV_][PRP_][DET_F_S_][N_F_S_]	1	3	8	12
651	[N_F_P_][PRP_][DET_F_S_][N_F_S_]	1	2		
665	[N_M_S_][N_M_S_]	2	6	3	47
669	[N_F_S_][PRP_][DET_F_S_][N_F_S_]	3	5		
679	[N_F_S_][ADV_]	1	4		
687	[N_M_S_][V_PR_3S_IND_VFIN_]	1	3	4	5
731	[N_M_S_][ADJ_M_S_][PRP_][DET_M_S_][N_M_S_]	2	3		

Figura 9 – Laudo de ressonância magnética antes do pré-processamento

Pequenos focos de hipersinal nas sequências com TR longo são vistos em núcleos da base bilateralmente (gliose? desmielinização?).  
 Presença de hipersinal nas sequências com TR longo em nervo óptico esquerdo.  
 Discreta assimetria de espessura dos nervos ópticos, maior à esquerda.  
 Ventrículos cerebrais, sulcos corticais, cisternas e fissuras dentro da normalidade.  
 Cerebelo e tronco cerebral sem alterações.  
 Ausência de coleções extra axiais.  
 Espessamento mucoso do seio esfenoidal e de células etmóides.  
 As sequências de angiorressonância não evidenciam alterações dignas de nota.

Figura 10 – Laudo de ressonância magnética depois do pré-processamento

pequenos focos de hipersinal em as sequências com tempo de repetição longo são vistos em núcleos de a base bilateralmente ( gliose? desmielinização? ) .  
 presença de hipersinal em as sequências com tempo de repetição longo em nervo óptico esquerdo . discreta assimetria de espessura de os nervos ópticos , maior a a esquerda . ventrículos cerebrais , sulcos corticais , cisternas e fissuras dentro de a normalidade . cerebelo e tronco cerebral sem alterações .  
 ausência de coleções extra axiais . espessamento mucoso de o seio esfenoidal e de células etmóides .  
 as sequências de angiorressonância não evidenciam alterações dignas de nota .

O texto nas Figuras 9 e 10 são dos laudos de Ressonância Magnética utilizados, antes e depois do pré-processamento do texto, respectivamente. O Quadro 8 mostra a estruturação do mesmo, sendo que na primeira coluna temos os AR, que representam o que foi encontrado, e a segunda coluna temos os TM, que representam onde o achado foi encontrado.

Esta estruturação foi obtida pela concatenação dos achados de mesmo elemento. Ou seja, AR ou TM contíguos ou intercalados. Por exemplo, o primeiro AR é composto das seguintes identificações: “focos de hipersinal”, “sequências com tempo” e “tempo de repetição longo”, sendo que as duas primeiras são contíguas e as duas últimas são intercaladas, pois tem a palavra “tempo” em comum.

Quadro 8 – Estruturação do laudo da Figura 9

AR	TM
focos de hipersinal sequências com tempo de repetição longo	
	núcleos de a base bilateralmente
gliose desmielinização	
presença de hipersinal sequências com tempo de repetição longo	
	nervo óptico esquerdo
assimetria de espessura de os nervos	
	nervos opticos
espessamento mucoso de o seio	
	seio esfenoidal
	espessamento mucoso
células etmóides	

Esta estruturação foi obtida pela concatenação dos achados de mesmo elemento. Ou seja, AR ou TM contíguos ou intercalados. Por exemplo, o primeiro AR é composto das seguintes identificações: “focos de hipersinal”, “sequências com tempo” e “tempo de repetição longo”, sendo que as duas primeiras são contíguas e as duas últimas são intercaladas, pois tem a palavra “tempo” em comum.

#### 4.2.2 Utilização do ESTEMED por profissionais de saúde

Na fase 3 foram criadas 307 regras por profissionais de saúde da equipe de teste, utilizando o ESTEMED. Destas, 105 não encontraram textos equivalentes na base de teste. Estas não serão portanto consideradas nos resultados, a não ser na Tabela 7, que apresenta a soma da quantidade de regras encontradas por grupo de 5 textos, visando mostrar o aumento ou diminuição da quantidade de regras criadas ao longo do processo.

Tabela 7 – Equipe de Teste - Quantidade de regras por grupo de 5 textos.

textos			Soma de regras
De	à		
1	5		40
6	10		34
11	15		72
16	20		19
21	25		31
26	30		34
31	35		17
36	40		20
41	45		19
46	50		21

Na Tabela 8 são apresentados os resultados correspondentes às regras agrupadas pelo número de termos das mesmas. Por exemplo, a regra “[N\_M\_S\_] [N\_M\_S\_]” tem dois termos, a regra “[N\_F\_S\_] [PRP\_] [N\_M\_S\_]” tem três termos, e assim por diante. Com esta tabela pode-se observar qual a quantidade de termos que geraram melhores resultados, em comparando com a quantidade de regras criadas com a respectiva quantidade de termos.

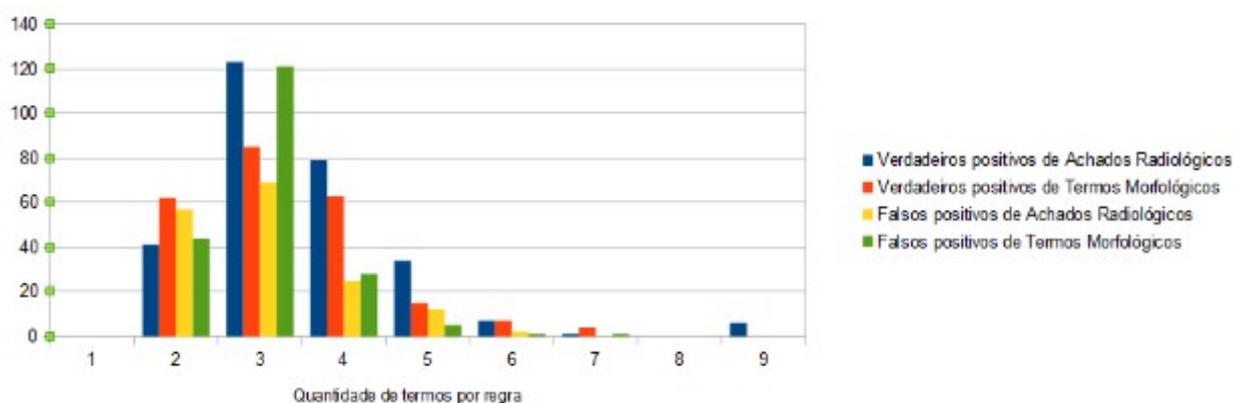
Tabela 8 - Equipe de Teste - Quantidade de regras por número de termos, com a quantidade de trechos identificados no texto e quantidade de VP e FP

Número de Termos por Regra	Total de Regras			Verdadeiros Positivos			Falsos Positivos			Total		
	AR	TM	T	AR	TM	T	AR	TM	T	AR	TM	T
2	7	11	18	41	62	103	57	44	101	98	106	204
3	37	29	66	123	85	208	69	121	190	192	206	396
4	44	16	60	79	63	142	25	28	53	104	91	195
5	25	11	36	34	15	49	12	5	17	46	20	66
6	3	8	11	7	7	14	2	1	3	9	8	17
7	4	4	8	1	4	5	0	1	1	1	5	6
9	1	0	1	6	0	6	0	0	0	6	0	6

A quantidade total de trechos identificados na base de teste foi 1996 sendo que houve 1744 casos de intersecções entre AR e TM, ou seja, partes de trechos identificados como AR e TM.

A Figura 11 mostra um gráfico que permite uma melhor comparação nos resultados obtidos pelo conjunto de regras, separadas pela quantidade de termos.

Figura 11 - Equipe de Teste - VP e FP por quantidade de termos das regras



A Tabela 9 apresenta os valores de VP, FP, FN e os valores de sensibilidade e precisão.

Tabela 9 – Equipe de Teste - Valores obtidos

<i>F-measure</i>	Precisão	Sensibilidade	VP	FP	FN
0,679487	0,5882	0,8042	530	371	129

Na Tabela 10 pode-se observar a variação na quantidade de identificações obtidas nos textos, pois ela apresenta a média e o desvio padrão dos VP e FP por texto considerando primeiramente a combinação dos resultados de AR, TM e a seguir considerando somente os AR e depois os TM.

Tabela 10 – Equipe de Teste - Média e desvio padrão de VP e FP por texto

	Média	Desvio Padrão	Média AR	Desvio Padrão AR	Média TM	Desvio Padrão TM
VP por Texto	10,62	9,62	6,47	5,53	5,62	4,60
FP por texto	9,28	7,26	5,16	3,84	5,13	4,12

A Tabela 11 mostra o resultado apenas das regras que foram consideradas de performance superior na base de teste. Esta seleção foi feita pegando as regras com maior número de VP e com até 3 FP. A Tabela 12 mostra o resultado apenas das regras que foram consideradas de performance inferior na base de teste. Estas foram selecionadas quando obtiveram mais resultados FP do que VP. Os resultados das Tabelas 11 e 12 permitem uma comparação em termos de proporção com os resultados obtidos pela equipe de desenvolvimento.

Tabela 11 – Equipe de Teste - Resultados das melhores regras

Id	Composição	Teste	
		VP	FP
1516	[PRP_][DET_M_S_][N_M_S_][ADJ_M_S_]	32	2
1703	[N_M_S_][PRP_][N_M_S_]	31	3
1613	[PRP_][N_F_S_][ADJ_F_S_][ADJ_F_S_]	18	1
2060	[N_M_S_][PRP_][DET_F_S_][N_F_S_][ADJ_F_S_]	16	2
1842	[N_M_S_][PRP_][N_F_S_]	14	2
1735	[N_M_P_][ADJ_M_P_][-PNT_NSEP]	13	1
1884	[-PNT_REL][N_F_S_]	13	1

Tabela 12 – Equipe de Teste - Resultados das piores regras

Id	Composição	Teste	
		VP	FP
1846	[N_F_P_][KC_][N_F_P_]	4	7
1782	[N_M_S_][PRP_][DET_M_P_][N_M_P_]	3	4
2162	[N_F_S_][KC_][N_F_S_]	3	6
1494	[PRP_][N_F_S_][PRP_][ADJ_F_S_]	1	2
1727	[ADV_][DET_F_S_][N_F_S_]	1	2
1818	[V_GER_][DET_M_S_][N_M_S_]	1	2

#### 4.3 MÉTODO PROPOSTO

Os passos deste método são descritos abaixo:

**Passo 1:** Randomizar o conjunto de laudos e separar um conjunto para treinamento e um para teste, colocando em uma tabela no formato Microsoft Excel.

**Passo 2:** Utilizar o ESTEMED para importar os conjuntos de teste e treinamento.

**Passo 3:** Utilizar a base de treinamento na tela de cadastramento de regras para cadastrar todas as regras para os itens de interesse. Esta criação de regras deve utilizar os seguintes princípios:

a) As regras que geram melhores resultados com relação a identificação de AR e TM são compostas por 3,4 ou 5 palavras. As regras com menos de 3 gerarão um número maior de FP e as com mais de 5 terão poucos VP.

b) Se um conjunto com mais de 6 palavras for de interesse, verificar mentalmente se outras situações se aplicam a mesma sequência de classificações morfológicas e caso não se aplique, verificar a possibilidade de quebrar em duas regras.

c) Quando uma regra for muito pequena, por exemplo, com duas palavras, deve-se verificar se não é possível agregar artigo ou preposição antes ou depois, e em último caso, agregar uma vírgula, se houver.

**Passo 4:** A cada 10 textos com regras criadas, aplicar estas regras com a opção executar do ESTEMED, selecionando a base de treinamento para verificar se as regras que estão sendo geradas estão dando o resultado esperado. Para realizar esta avaliação deve ser criada uma lista com a quantidade de VP e FP de cada regra. As que obtiverem maior quantidade de FP ou maior quantidade de FP do que VP devem ser analisadas para identificar possibilidades de melhoramentos. Nas regras com mais de 4 termos que tenham menos de 3 VP deve ser feita uma tentativa de separação dos termos para obter-se duas regras com menos termos. Nas regras com dois ou três termos que tenham as maiores quantidades de FP (aproximadamente os 10% que encabecem a lista) deve ser procurada uma palavra ou uma pontuação para ser incluída na regra e torná-la mais específica.<sup>1</sup>

---

<sup>1</sup>Os passos 4, 8, 9, 10 e 11 são realizado utilizando acesso direto à base de dados, pois o ESTEMED não possui ainda estas funcionalidades.

**Passo 5:** A partir da tela de resultado, verificar quais os trechos que devem ser encontrados e que não foram e passar para a tela de cadastramento para criar as regras correspondentes.

**Passo 6:** Repetir os passos três, quatro e cinco para todos os laudos de treinamento.

**Passo 7:** Executar as regras na base de treinamento e observar os resultados. Utilizar a tela de resultados para marcar os FN e marcar nas observações o que for FP.

**Passo 8:** Para todos os FP e FN relacionados a erros de classificação morfológica da palavra, feita pelo CoGrOO, ocasionado pela limitação do *corpus* anotado, criar registro na tabela que corrige estas situações.

**Passo 9:** Identificar todos os resultados cuja palavra anterior é um artigo e verificar se a presença do artigo identifica se é FP ou VP. Se for o caso, eliminar os resultados com o artigo que identifica os FP.

**Passo 10:** Quando dois resultados são de mesmo elemento (AR ou TM) e um está completamente contido no outro, o primeiro deve ser eliminado.

**Passo 11:** Depois de todos os ajustes realizados executar as regras na base de teste, executar na base de teste os passo 8, 9 e 10 e anotar todos os VP, FP e FN para calcular especificidade e sensibilidade e *f-measure*. Serão considerados como VP quando uma regra encontrar um texto que corresponde ao mesmo elemento da regra (AR ou TM) ou que tenha parte dele que corresponda. Caso contrário será considerado FP. Qualquer palavra ou conjunto de palavras que faça parte de um AT ou TM e que não tenha sido encontrado por nenhuma regra será considerado FN. As fórmulas utilizadas para calcular as proporções são:

$$precisão = \frac{VP}{(VP+FP)} \quad (1)$$

$$sensibilidade = \frac{VP}{(VP+FN)} \quad (2)$$

$$f\text{-measure} = 2 * \frac{\text{precisão} * \text{sensibilidade}}{(\text{precisão} + \text{sensibilidade})} \quad (3)$$

Vale ressaltar que a execução do passo 11 pode ocorrer ou não, dependendo se o interesse é em somente obter as identificações de AR e TM ou se também se quer saber quais os valores das proporções obtidos.

Vale ressaltar que para funcionamento, o ESTEMED e o banco de dados devem estar instalados e deve-se ter as tabelas auxiliares devidamente preenchidas, salientando a tabela “elementos” que identifica quais os tipos de informação que serão recuperadas.

## 5 DISCUSSÃO

Neste capítulo serão discutidos o método que utiliza o ESTEMED, o processo da criação das regras e utilização do método, a comparação dos valores de precisão e sensibilidade e trabalhos futuros.

### 5.1 O MÉTODO

O método proposto neste trabalho se baseia nas experiências anteriores de Bulegon (2011) e Oliveira(2014). Nestes trabalhos os objetivos foram bastante diferentes. No primeiro foi uma recuperação de informação (diagnóstico), no segundo a identificação da presença de uma informação (continuidade de cuidado) ambos utilizando *corpus* de resumo de alta.

Comparando com os valores obtidos por Oliveira (2014), observamos que foram superiores aos do presente trabalho, apesar de utilizar as regras de classificação morfológica e a maior parte dos processos são iguais. Parte da explicação para este fato é a diversidade de estruturas de frases que se encontram em cada uma das situações e a quantidade de ocorrências em cada texto. Esta diversidade foi ainda mais relevante devido ao fato dos laudos que pertencem ao período inicial corresponderem a uma fase de evolução intensa, pois o equipamento de ressonância magnética havia sido recentemente adquirido.

As continuidades de tratamento têm poucas ocorrências nos textos e não ocorrem em número significativo de textos. Já os AR e os TM foram recuperados de laudos radiológicos, que são textos extremamente concisos e cada frase se refere a pelo menos um AR e um TM, a não ser nas frases que indicam a normalidade de alguma estrutura morfológica.

Assim sendo, com a utilização do ESTEMED recuperou-se em média 12,14 informações por laudo, como indica a Tabela 4 e a diversidade de estruturas é significativamente maior, pois o trabalho de Oliveira (2014) utilizou 30 regras, enquanto com o ESTEMED foram criadas 182 regras. Também o fato de 45 regras

criadas na base de treinamento não terem sido encontradas na base de teste indica a diversidade de estruturas.

Para a característica específica dos textos dos laudos processados foi percebido que o método de Oliveira (2014) não era totalmente adequado. A diversidade, mais uma vez, é a provável causa de pequenas diferenças como vírgulas ou artigos em determinados textos serem importantes para a identificação dos AR e TM e portanto os métodos que removiam as vírgulas e artigos tiveram que ser eliminados. Houve também a necessidade da criação de um processo de identificação de frases que descrevem situações de normalidade devido ao estilo de escrita dos laudos, que se baseiam em frases prontas para diversas situações mais comuns, e que os radiologistas utilizam sem alterações para descrever as situações normais dependendo da causa da solicitação do exame.

Na fase inicial da pesquisa foi possível observar as dificuldades em se determinar as regras baseadas em classes morfológicas das palavras. Sem um método pré-determinado e sem indicações de como construir as regras, houve perda de tempo com a eliminação de regras inadequadas e criação de regras substitutas. Na primeira fase, etapa 1.5, já com uma parte de método sendo utilizado e com algum conhecimento adquirido no que tange a formação de regras com mais chances de dar bons resultados, o processo correu mais naturalmente e com maior produtividade. Aspectos como esses não são citados nas pesquisas utilizadas como referencial, exceto Oliveira (2014) que aborda o assunto, em um trabalho que foi realizado no mesmo grupo de pesquisa.

Já na fase 4, quando os testes foram realizados por médicos e enfermeiras, como o ESTEMED foi desenvolvido tendo em mente a utilização por profissionais de saúde, foi observada a facilidade na utilização pelos mesmos. Esta facilidade ficou evidenciada devido ao pouco tempo de treinamento e ao fato de que as dúvidas da equipe de teste basicamente se restringiram à conveniência ou não na criação da regra, e não de como criá-la no sistema. A maior dificuldade da equipe de teste foi relacionada às dúvidas no método. Em diversas situações as recomendações não foram suficientemente específicas e a conveniência de criar uma regra com diferentes características deixaram os testadores indecisos. Pode-se dizer então que a experiência direta no processo não pode ser substituída pelas indicações

presentes no método, mas a curva de aprendizado se mostrou bastante favorável, pois a criação das regras foi realizada em torno de 1:30 horas em média, contando com o tempo para a informar o método que devia ser utilizado.

Vale ressaltar que diferentemente dos trabalhos de Bulegon (2011) e Oliveira (2014), nesta pesquisa o sistema foi utilizado por profissionais de saúde. Sendo que os médicos e enfermeiras não tiveram dificuldades em utilizar o ESTEMED. Além disso, foi fundamental a criação de regras pelos profissionais de saúde, uma vez que esta tarefa deve ser realizada por quem tem conhecimento das necessidades e está habituado a elaborar laudos de imagens. Nos testes iniciais as regras foram criadas por uma acadêmica de medicina, no entanto, na etapa 2.1 “Criação das regras” pela equipe de desenvolvimento não houve disponibilidade deste recurso, e as regras foram criadas pelo analista de sistema, que devido a dificuldades nesta atividade teve que recorrer à marcação prévia em papel por parte de profissionais de saúde. Isto reforça a importância da ESTEMED possibilitar e facilitar a participação de médicos na criação de regras.

Os *corpora* anotados utilizados em outras pesquisas são bem adaptados previamente aos textos das narrativas que compõe o corpus que se pretende pesquisar, ou ao longo da pesquisa é realizada a adaptação necessária. Dependendo do idioma, já existem *copus* da área médica, como o FraMed, que é em alemão (Gerstmair *et al.*, 2012). Como não havia disponível um *corpus* anotado para laudos radiológicos de ressonância magnética em português, levantamos a hipótese de que a utilização de *corpus* anotado composto por resumos de alta poderia ser utilizado com sucesso em laudos radiológicos, o que não se confirmou, pois foi necessário a alteração no processo para corrigir a classificação morfológica de palavras que estavam sendo classificadas erroneamente pelo CoGrOO. Foram 120 nos testes com a equipe de desenvolvimento e mais 15 com a equipe de teste.

Como foi citado anteriormente, tanto os métodos de estatística como os de regras, muitas vezes também utilizam dicionários de forma combinada, como podemos encontrar em (Friedman, 2000). Fica claro neste artigo a dificuldade em termos de quantidade de recursos humanos para se criar ou aperfeiçoar um dicionário para cada área. Devido a esta dificuldade, a proposta do presente trabalho foi criar um sistema que não dependa de dicionário.

## 5.2 O PROCESSO DA CRIAÇÃO DAS REGRAS E A UTILIZAÇÃO DO MÉTODO NO ESTEMED

Como afirma Lacson *et al.* (2012), o desenvolvimento com direcionamento para a facilidade de uso permite a utilização do sistema de forma que usuários sem a formação técnica específica usem-no sem maiores dificuldades. Da equipe de teste, 6 membros utilizaram o ESTEMED, sendo três enfermeiras e três médicos e não tiveram dificuldades na utilização do sistema, mesmo com um treinamento extremamente rápido (em torno de 30 minutos). Assim, a hipótese de que a tarefa de criação de regras de classificação morfológicas poderia ser simplificada a ponto de profissionais de saúde poderem assumi-la com pouco treinamento foi confirmada tendo como textos bases os laudos de Ressonância Magnética de Encéfalo. Todavia, devemos acrescentar que houve uma relativa dificuldade e um número expressivo de dúvidas quanto a conveniência ou não na criação de regras e algumas situações específicas. Isso se deveu principalmente a diversidade de estilos de escrita, a complexidade dos textos e em algumas situações a própria complexidade do AR ou TM em questão.

Honorato *et al.* (2008) utilizaram PLN para facilitar o trabalho de especialistas na formação de um dicionário com todas as frases presentes nos laudos, e o trabalho de estruturação foi a identificação binária se uma frase está presente em um laudo ou não. Neste caso, foram utilizados laudos de Endoscopia Digestiva, que já são separados conforme a parte do sistema digestivo que referencia, o que representa uma certa estruturação. Também devemos considerar que o vocabulário é limitado e as frases são assertivas simples. Nos laudos utilizados no ESTEMED não há qualquer tipo de estrutura ou separação em tópicos, a terminologia utilizada é extensa e as frases não são padronizadas a não ser as utilizadas para identificar a ausência de AR. Estes fatores dificultam a obtenção de melhores proporções entre VP, FN e FP nos resultados dos experimentos.

Outro ponto importante na utilização do PLN é o tamanho dos *corpora* de treinamento e teste. Como podemos observar no Quadro 2, a quantidade de textos utilizados varia significativamente entre os artigos. Dos 23.000 laudos disponíveis, o ideal seria a utilização de todos, no entanto, a quantidade de recursos humanos necessária para tal seria inviável no contexto desta pesquisa. Assim trabalhamos com 300 laudos para a equipe de desenvolvimento, sendo que 50 para testes iniciais durante a prototipação, 250 para treinamento (50) e teste (200), e 100 para a equipe de teste, sendo 50 para treinamento e 50 para teste. A comparação com outros trabalhos também deve considerar que diversos deles tiveram quantidades menores de identificações por texto, muitas vezes menos e uma em média.

Os resultados obtidos foram abaixo do esperado, tendo em vista que foi utilizada a mesma técnica de Oliveira (2014). No entanto, deve-se considerar que os recursos para criação de regras foram limitados, e que um dos objetivos da técnica utilizada é exatamente este, ou seja, utilizar poucos recursos. Como explicam Dublin *et al.* (2013), a limitação nos recursos para análise dos textos influencia diretamente a expectativa dos percentuais de acerto e erro da técnica.

Com base nos testes iniciais realizados na fase de desenvolvimento do ESTEMED, a quantidade de regras novas estava diminuindo significativamente à medida que os 90% da base de treinamento eram atingidos quando foram usados 50 laudos, o que pareceu confirmar que a quantidade de laudos para esta base estava adequada. No entanto, a Tabela 1 mostra que os testes subsequentes não confirmaram esta situação, pois nos últimos 5 laudos foi criada uma quantidade de regras comparável com a dos 5 primeiros laudos. Isso indica que a base de treinamento foi provavelmente insuficiente para a variedade de estilos utilizados nos textos dos laudos.

Há que se considerar também que os resultados indicam uma base de treinamento insuficiente. Não só o número de regras não diminuiu nos laudos finais, como também não houve a mesma proporção entre os resultados da base de treinamento e os de teste nem das regras de melhor performance nem nas de pior, conforme podemos observar nas Tabelas 1, 5 e 6.

Pode-se também comparar o presente trabalho com outros que utilizam técnicas diferentes, e que da mesma forma recuperam informação genérica, como a

que foi feita (Lacson *et al.*, 2012; Gerstmair *et al.*, 2012), Observa-se neste último, que apesar de demonstrar sensibilidade e precisão superiores, só conseguiu melhores resultados com a extensão do dicionário que foi utilizado originalmente. Apesar dos objetivos de cada trabalho terem similaridade técnica, pois recuperam termos importantes mas não específicos de um texto médico, a forma de cálculo de precisão e sensibilidade é diferente, dificultando qualquer comparação. Também devemos considerar que no método proposto o objetivo era ter uma sensibilidade o melhor possível, de forma que uma quantidade relativamente grande de FP encontrados pela equipe de desenvolvimento na base de teste (em média 4,85 por texto) foi propositadamente tolerada.

A comparação entre os testes realizados pela equipe de desenvolvimento e pela equipe de testes deve ser realizada com alguns cuidados, pois o método utilizado pela equipe de teste foi bastante simplificado, tendo em vista que o objetivo principal era identificar as dificuldades na utilização do ESTEMED. Como era de se esperar, sem ter passado por ciclos de revisão e aperfeiçoamento das regras, e equipe de testes obteve uma quantidade proporcionalmente bem maior de FP e de intersecções entre AR e TM. No entanto, a quantidade de FN foi aproximadamente de dentro da proporção, dando um resultado de sensibilidade próximo ao dos testes da equipe de desenvolvimento.

### 5.3 TRABALHOS FUTUROS

Consideramos que os resultados obtidos com o ESTEMED podem ser utilizados em atividades de mineração de textos, ajudando a identificação dos *n-grams* mais importantes. Da mesma forma, poderia ser o primeiro passo na confecção de uma ontologia ou dicionário para a área que está sendo pesquisada.

A técnica de regras de classificação morfológica pode ser aprimorada neste contexto com:

a) Contar com a utilização de sub-regras, com o objetivo de confirmar ou descartar uma ocorrência de uma regra principal. Estas sub-regras deveriam se referir não somente com palavras próximas como também com ocorrências em

distâncias pré-determinadas. Este tópico foi parcialmente testado mas é necessário que seja aprofundado.

b) Realizar estudo estatístico dos *n-grams* na base completa para ajudar a formação de uma base de treinamento o mais diversa possível, facilitando a tarefa de criação das regras.

c) Permitir que o usuário determine na formação das regras se a palavra será representada pela sua classificação morfológica, como é atualmente, ou pela função sintática, ou pela raiz da palavra, ou pela palavra inteira.

O ESTEMED foi desenvolvido com funcionalidades básicas, que podem ser significativamente ampliadas, principalmente em telas de resultados que permitam uma melhor compreensão do desempenho de cada regra, e possivelmente sugerir alterações nas mesmas de forma a atingir melhores porções entre VP, FN e FP.

A utilização do ESTEMED poderia ser integrada com o Morphosaurus, para facilitar a tradução dos termos ou com a nomenclatura SNOMED traduzida para o português, visando a estruturação dos textos baseada em uma terminologia padronizada.

O método e o sistema propostos também podem ser testados primeiramente em laudos do mesmo tipo em diferentes instituições, e posteriormente em outros tipos de laudos.

As dificuldades encontradas na classificação morfológica das palavras nos laudos indicam a conveniência de se desenvolver um corpus anotado para esta área.

## 6 CONCLUSÃO

A utilização do ESTEMED pela equipe de teste comprovou a hipótese que esta técnica e método podem ser utilizados por profissionais de saúde com pouco treinamento e suporte. Fica claro, no entanto, a necessidade de evolução do sistema, principalmente com a criação de telas que mostrem os resultados de cada regra criada, para facilitar a identificação e consequente eliminação das regras com desempenho que o usuário considere inadequado.

A tentativa de utilização de corpus anotado formado por resumos de alta não foi bem sucedida, tendo em vista que a terminologia dos laudos radiológicos contém uma quantidade significativa de termos que não estão presentes nos resumos de alta. Esta situação fica clara com a observação dos 135 termos colocados em tabela do banco de dados para que o ESTEMED corrigisse os termos classificados erroneamente.

A evolução do método utilizado por Bulegon (2011) e por Oliveira (2014) permitiu encontrar novos usos e identificar limites da técnica para a continuidade das pesquisas relativas a utilização de PLN com regras de classificação morfológica.

A com melhorias a serem realizadas na técnica, no método e no ESTEMED, há um potencial de contribuição para a evolução do PLN na área médica e para a difusão na utilização da mesma, possibilitando uma melhoria na qualidade dos serviços relacionados.

## REFERÊNCIAS

AMBERT, K. H.; COHEN, A. M. A system for classifying disease comorbidity status from medical discharge summaries using automated hotspot and negated concept detection. **Journal of the American Medical Informatics Association : JAMIA**, v. 16, n. 4, p. 590–5, 2009.

BORKAR, A. T.; NIKOSE, A. A.; CHATURVEDI, B. Knowledge Extraction From Radiology Report Using Text Mining. **International Journal of Engineering Research & Technology (IJERT)**, v. 2, n. 8, p. 23–28, 2013.

BUI, Q.-C. *et al.* Extracting causal relations on HIV drug resistance from literature. **BMC bioinformatics**, v. 11, p. 101, jan. 2010.

BULEGON, H. **Identificação de diagnósticos contidos em narrativas clínicas e mapeamento para a classificação internacional de doenças**. 2011. 101 f. Dissertação - Pontifícia Universidade Católica do Paraná. Curitiba, 28 de fev., 2011.

CHAPMAN, W. W.; COHEN, K. B. Current issues in biomedical text mining and natural language processing. **Journal of biomedical informatics**, v. 42, n. 5, p. 757–9, out. 2009.

CHILDS, L. C. *et al.* Description of a rule-based system for the i2b2 challenge in natural language processing for clinical data. **Journal of the American Medical Informatics Association : JAMIA**, v. 16, n. 4, p. 571–5, 2009.

CLIMENSON, W. D. Language-Analysis Problems in the Computer Processing of Natural Text. **ENGINEERING WRITING AND SPEECH**, v. EWS-6, n. 2, p. 72–78, 1963.

DAI, H.-J. *et al.* Coreference resolution of medical concepts in discharge summaries by exploiting contextual information. **Journal of the American Medical Informatics Association : JAMIA**, v. 19, n. 5, p. 888–96, 2012.

DO, B. H. *et al.* Automatic Retrieval of Bone Fracture Knowledge Using Natural Language Processing. **Journal of digital imaging : the official journal of the Society for Computer Applications in Radiology**, 2 out. 2012.

DUBLIN, S. *et al.* Natural Language Processing to identify pneumonia from radiology reports. **Pharmacoepidemiology and Drug Safety**, 2013.

DUTTA, S. *et al.* Automated Detection Using Natural Language Processing of Radiologists Recommendations for Additional Imaging of Incidental Findings. **Annals of emergency medicine**, p. 1–8, 30 mar. 2013.

FENELON, S. Aspectos ético-legais em Imaginologia. **Radiologia Brasileira**, v. 36, n. 1, p. III–VI, 2003.

FRIEDMAN, C. A broad-coverage natural language processing system. **Proceedings / AMIA ... Annual Symposium. AMIA Symposium**, p. 270–4, jan. 2000.

GERSTMAIR, A. *et al.* Intelligent image retrieval based on radiology reports. **European radiology**, 4 ago. 2012.

HONORATO, D. D. F. *et al.* Construction of an Attribute-Value Representation for Semi-structured Medical Findings Knowledge Extraction. **CLEI electronic journal**, v. 11, n. 3, p. 1–12, 2008.

KAHN, C.; LANGLOTZ, C.; BURNSIDE, E. Toward Best Practices in Radiology Reporting1. **Radiology**, v. 252, n. 3, p. 1–5, 2009.

LACSON, R. *et al.* Information from Searching Content with an Ontology-Utilizing Toolkit (iSCOUT). **Journal of digital imaging**, v. 25, n. 4, p. 512–9, ago. 2012.

LI, S.-T.; CHEN, C.-C.; HUANG, F. Conceptual-driven classification for coding advise in health insurance reimbursement. **Artificial intelligence in medicine**, v. 51, n. 1, p. 27–41, jan. 2011.

OLIVEIRA, L. **Identificação de continuidade de tratamento em sumários de alta hospitalar**. 2014. 66 f. Dissertação - Pontifícia Universidade Católica do Paraná. Curitiba. 2014.

PATTERSON, O.; HURDLE, J. F. Document clustering of clinical narratives: a systematic study of clinical sublanguages. **AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium**, v. 2011, p. 1099–107, jan. 2011.

PERCHA, B. *et al.* Automatic classification of mammography reports by BI-RADS breast tissue composition class. **Journal of the American Medical Informatics Association : JAMIA**, v. 19, n. 5, p. 913–6, 2012.

POOL, F.; GOERGEN, S. Quality of the Written Radiology Report : A Review of the Literature. **Journal of the American College of Radiology (JACR)**, v. 7, n. 8, p. 634–643, 2010.

PRASAD, A.; RAMAKRISHNA, S. Extraction of Radiology Reports using Text mining. **International Journal on Computer Science and Engineering**, v. 02, n. 05, p. 1558–1562, 2010.

REINER, B. I.; KNIGHT, N.; SIEGEL, E. L. Radiology reporting, past, present, and future: the radiologist's perspective. **Journal of the American College of Radiology : JACR**, v. 4, n. 5, p. 313–9, maio. 2007.

REZENDE, S. O. Mineração de Texto. *In: Sistemas Inteligentes: fundamentos e aplicações*. Barueri: Manole, 2005. p. 337–370.

SALMASIAN, H.; FREEDBERG, D. An automated tool for detecting medication overuse based on the electronic health records. **pharmacoepidemiology and drug safety**, 2012.

SCHULZ, S. *et al.* **MORPHOSAURUS - Crosslingual Medical Text Retrieval by Subword Indexing** AMIA Annual Symposium proceedings. **Anais...**2002

SERAPIÃO, P. *et al.* Construção de um índice de informação da prática clínica em Radiologia e Diagnóstico por Imagem baseada em mineração de texto e tesauro \*. **Radiologia Brasileira**, v. 46, n. 5, p. 290–298, 2013.

STAMATATOS, E.; FAKOTAKIS, N.; KOKKINAKIS, G. Automatic text categorization. **Computational Linguistics**, v. 26, n. 4, p. 471–495, 2005.

WANG, X. *et al.* Active computerized pharmacovigilance using natural language processing, statistics, and electronic health records: a feasibility study. **Journal of the American Medical Informatics Association : JAMIA**, v. 16, n. 3, p. 328–37, 2009.

YANG, H. *et al.* A text mining approach to the prediction of disease status from clinical discharge summaries. **Journal of the American Medical Informatics Association : JAMIA**, v. 16, n. 4, p. 596–600, 2009.

YANG, H. Automatic extraction of medication information from medical discharge summaries. **Journal of the American Medical Informatics Association : JAMIA**, v. 17, n. 5, p. 545–8, 2010.

YETISGEN-YILDIZ, M. *et al.* A text processing pipeline to extract recommendations from radiology reports. **Journal of biomedical informatics**, v. 46, n. 2, p. 354–362, 24 jan. 2013.