



Universidade de Brasília

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

**Reconhecimento de Entidades Nomeadas na base de  
notificações de eventos adversos e queixas técnicas de  
dispositivos médicos no Brasil**

Gustavo Cunha Garcia

Dissertação apresentada como requisito parcial para conclusão do  
Mestrado Profissional em Computação Aplicada

Orientador  
Prof. Dr. Marcelo Ladeira

Brasília  
2021

## **Ficha Catalográfica de Teses e Dissertações**

Esta página existe apenas para indicar onde a ficha catalográfica gerada para dissertações de mestrado e teses de doutorado defendidas na UnB. A Biblioteca Central é responsável pela ficha, mais informações nos sítios:

<http://www.bce.unb.br>

<http://www.bce.unb.br/elaboracao-de-fichas-catalograficas-de-teses-e-dissertacoes>

**Esta página não deve ser incluída na versão final do texto.**



Universidade de Brasília

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

# **Reconhecimento de Entidades Nomeadas na base de notificações de eventos adversos e queixas técnicas de dispositivos médicos no Brasil**

Gustavo Cunha Garcia

Dissertação apresentada como requisito parcial para conclusão do  
Mestrado Profissional em Computação Aplicada

Prof. Dr. Marcelo Ladeira (Orientador)  
CIC/UnB

Prof.a Dr.a Nádia Félix Felipe Da Silva    Prof. Dr. Carlos Henrique Reis Esselin Rassi  
Instituto de Informática/UFG                      Faculdade de Medicina/UnB

Prof. Dr. Luís Paulo Faina Garcia  
CIC/UnB

Prof. Dr. Marcelo Ladeira  
Coordenador do Programa de Pós-graduação em Computação Aplicada

Brasília, 31 de agosto de 2021

# Dedicatória

Dedico esse trabalho aos meus pais, Neuma Cunha Garcia e Antônio Carlos de Almeida Garcia, que se sacrificaram para que eu pudesse chegar até aqui. Dedico também a minha esposa Renata C. de Carvalho Garcia e meus filhos Vinícius de Carvalho Garcia e Guilherme de Carvalho Garcia, por toda compreensão e incentivo nesses anos.

# Agradecimentos

Meus agradecimentos aos meus pais, minha irmã querida, minha esposa e meus filhos.

Gostaria de agradecer também ao meu orientador Marcelo Ladeira por toda a paciência, pelos encontros semanais, por toda a atenção, pelas palavras de incentivo, pela compreensão nos momentos difíceis, pelos ensinamentos e por ter aceitado o desafio de me orientar no meio do caminho.

Meus agradecimentos também aos meus companheiros de PPCA/UNB, que muito me ajudaram a enfrentar os desafios ao longo das matérias.

Agradeço também aos companheiros de ANVISA, da GGREG e, principalmente, aos colegas da GECOR, que não apenas compreenderam meu momento, como também pelas palavras de incentivo.

# Resumo

Os dispositivos médicos estão cada dia mais presentes no cotidiano das pessoas. Por se tratar de produtos que lidam diretamente com a saúde das pessoas, uma falha pode gerar graves consequências à saúde. Estudos pré mercado possuem uma série de limitações, o que torna o monitoramento pós mercado essencial, pois uma ação tempestiva pode evitar ainda maiores danos à população. Nesse sentido, o presente estudo tem como objetivo desenvolver uma solução baseada em aprendizado de máquina, utilizando uma técnica de mineração de texto conhecida como Reconhecimento de Entidade Nomeadas - NER. Após uma série de experimentos, o modelo BiLSTM + CRF, com a utilização de *word embedding* pré-treinadas e uma base de dados anotada com o padrão BIO alcançaram uma macro *F1-score* de 87.75%. Após a aplicação do modelo a toda base de dados, o modelo demonstrou boa capacidade em encontrar novas entidades na base de teste, que não estavam presentes na base de treino. As informações geradas pelo modelo permitiram a obtenção de informações importantes para a tomada de decisão da Agência, favorecendo uma intervenção tempestiva, o que evitaria maiores danos à saúde da população que seria exposta ao dispositivos médicos.

**Palavras-chave:** reconhecimento de entidade nomeada, dispositivos médicos, eventos adversos, queixas técnicas

# Abstract

Medical devices are increasingly present in people's daily lives. They are products that deal directly with people's health, and a failure can have serious health consequences. Pre-market studies have a number of limitations, which makes post-market surveillance essential, because timely action can prevent even greater damage to the population. In this sense, this study aims to develop a solution based on machine learning, using a text mining approach known as Named Entities Recognition - NER. After a series of experiments, the BiLSTM + CRF model, using pre-trained word embedding and an annotated database with the BIOES-style pattern, reached a macro F1-score of 87.75%. After applying the model to the entire database, the model demonstrated good ability to find new entities in the test database that were not present in the training database. The information generated by the model allowed obtaining important information for the Agency's decision-making, favoring a timely intervention, which would avoid further damage to the health of the population that would be exposed to medical devices.

**Keywords:** named entity recognition, medical devices, adverse events, technical complaints

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Definição do Problema . . . . .	2
1.2	Justificativa do Tema . . . . .	5
1.3	Objetivos . . . . .	6
1.3.1	Objetivo Geral . . . . .	6
1.3.2	Objetivos Específicos . . . . .	7
1.4	Contribuições Esperadas . . . . .	7
1.5	Estrutura do Trabalho . . . . .	8
<b>2</b>	<b>Revisão de Estado da Arte</b>	<b>9</b>
2.1	Preparação da Pesquisa . . . . .	9
2.1.1	Trabalhos Correlatos - aplicação de NER em Tecnovigilância . . . . .	12
2.2	Aplicações do NER . . . . .	14
2.3	Métodos de NER . . . . .	15
2.4	Fatores que influenciam os resultados em NER . . . . .	18
2.4.1	Língua . . . . .	18
2.4.2	<i>Features</i> . . . . .	19
2.4.3	Representação de Segmento . . . . .	20
2.5	Avaliação . . . . .	20
<b>3</b>	<b>Fundamentação Teórica</b>	<b>23</b>
3.1	CRISP-DM . . . . .	23
3.2	Processamento de Linguagem Natural - NLP . . . . .	25
3.2.1	Extração de Informação . . . . .	26
3.2.2	Reconhecimento de Entidades Nomeadas - NER . . . . .	26
3.3	Avaliação do modelo proposto . . . . .	28
3.3.1	<i>Precision</i> . . . . .	29
3.3.2	<i>Recall</i> . . . . .	29
3.3.3	<i>F1-Score</i> . . . . .	29



3.4	<i>Conditional Random Fields - CRF</i> . . . . .	31
3.5	<i>Long-Short Term Memory</i> . . . . .	32
<b>4</b>	<b>Experimentos</b>	<b>33</b>
4.1	Metodologia de Pesquisa . . . . .	33
4.2	Entendimento do negócio . . . . .	34
4.3	Entendimento dos dados . . . . .	37
4.3.1	Estrutura dos formulários . . . . .	37
4.3.2	Estatísticas das notificações . . . . .	39
4.4	Preparação dos dados . . . . .	42
4.4.1	Dados selecionados . . . . .	42
4.4.2	Definição das classes e tags das Entidades Nomeadas . . . . .	43
4.4.3	Formação do Corpus . . . . .	44
4.4.4	Processo de Anotação . . . . .	46
4.5	Modelagem . . . . .	47
4.5.1	Atributos ou <i>features</i> . . . . .	48
4.5.2	<i>Conditional Random Fields - CRF</i> . . . . .	49
4.5.3	Arquitetura BiLSTM + CRF . . . . .	50
<b>5</b>	<b>Resultados e Análises</b>	<b>53</b>
5.1	Escolha do modelo e <i>features</i> . . . . .	53
5.1.1	Escolha das <i>features</i> e o impacto do <i>transfer learning</i> . . . . .	54
5.1.2	BiLSTM + CRF 50 épocas vs BiLSTM + CRF 100 épocas . . . . .	54
5.1.3	BiLSTM + CRF vs BiLSTM . . . . .	55
5.1.4	BiLSTM + CRF vs CRF . . . . .	55
5.2	Base em Português vs base traduzida para o Inglês . . . . .	56
5.3	Representação de segmento . . . . .	58
5.4	Aplicação do Modelo . . . . .	60
5.4.1	Detecção de alterações nas frequências de Queixas Técnicas de Cateter . . . . .	64
<b>6</b>	<b>Considerações Finais</b>	<b>67</b>
	Referências	69
	Anexo	75
<b>I</b>	<b>Formulário para Notificação de Evento Adverso associado ao uso de Artigo Médico-Hospitalar</b>	<b>76</b>

II	Formulário para Notificação de Queixa Técnica de Artigo Médico-Hospitalar	96
III	Formulário para Notificação de Evento Adverso Associado ao uso de Equipamento Médico-hospitalar	111
IV	Formulário para Notificação de Queixa Técnica de Equipamento Médico-Hospitalar	130

# Lista de Figuras

1.1	Quantidade de notificação ao ano. . . . .	4
2.1	Quantidade de publicações ao ano na base <i>Scopus</i> . . . . .	12
3.1	Fases do CRISP-DM. . . . .	24
3.2	Portões que regulam a passagem de informações. . . . .	32
4.1	Histograma quantidade de caracteres das notificações de QT e EA. . . . .	41
4.2	Modelo BiLSTM + CRF. . . . .	51
5.1	Painel de BI com os resultados da aplicação do modelo em toda a base de notificações de cateter. . . . .	63

# Lista de Tabelas

2.1	Quantidade de artigos com NER aplicado à tecnovigilância . . . . .	10
2.2	Quantidade de artigos com NER aplicado à farmacovigilância . . . . .	11
2.3	Quantidade de artigos sobre NER em 2020 e 2021 . . . . .	11
2.4	Quantidade de revisões e survey sobre NER entre 2020 e 2021 . . . . .	11
2.5	Algumas Representações de Segmento disponíveis na literatura . . . . .	20
3.1	Matriz de confusão . . . . .	28
4.1	Quantidade de campos por formulário . . . . .	38
4.2	Tipo de notificação . . . . .	39
4.3	Percentual de notificações por classe de produto . . . . .	39
4.4	Notificações por gravidade . . . . .	39
4.5	Notificações por tipo de gravidade . . . . .	40
4.6	Os 10 produtos com mais notificações . . . . .	40
4.7	As 10 classes de notificantes que mais notificaram . . . . .	41
4.8	Status das notificações em investigação . . . . .	42
4.9	Tags e significados das Entidades Nomeadas . . . . .	46
4.10	Tags e significados das Entidades Nomeadas . . . . .	47
4.11	Estatísticas sobre as anotações realizadas no corpus . . . . .	47
4.12	Atributos modelo base . . . . .	49
5.1	Comparação BiLSTM + CRF sem e com word embeddings pré-treinadas . . . . .	54
5.2	Comparação BiLSTM+CRF com treinamento de 50 e 100 épocas . . . . .	55
5.3	Comparação BiLSTM+CRF com BiLSTM . . . . .	55
5.4	Comparação BiLSTM+CRF com CRF para 20 amostras . . . . .	56
5.5	Comparação BiLSTM+CRF com CRF para 17 amostras . . . . .	56
5.6	Base português vs base traduzida para o inglês sem pré-treinamento . . . . .	57
5.7	Base português vs base traduzida para o inglês com pré-treinamento . . . . .	58
5.8	BIO vs BILOU . . . . .	59
5.9	IOE vs BILOU . . . . .	59

5.10 BIO vs IOE . . . . .	59
5.11 Estatísticas sobre a base de treino e teste . . . . .	60
5.12 Estatísticas sobre as Entidades classificadas pelo modelo . . . . .	61
5.13 Estatísticas sobre as classificações das entidades . . . . .	62

# Capítulo 1

## Introdução

O mercado de saúde mundial está se tornando complexo. A necessidade de se prevenir doenças, realizar o diagnóstico precoce de enfermidades e obter acesso a um tratamento eficaz vem demandando dispositivos médicos mais apropriados, seguros e efetivos. Os dispositivos médicos são cruciais na prevenção, diagnóstico, no tratamento de doenças e na reabilitação dos pacientes. Esses produtos estão cada vez mais presentes na vida das pessoas[1]. Os avanços tecnológicos tornaram os produtos médicos mais acessíveis e presentes no cotidiano. Seja nos consultórios, clínicas, hospitais ou mesmo nas residências, eles vêm auxiliando no cuidado à saúde e melhorando a qualidade de vida de seus usuários. Os termômetros digitais e infravermelhos, que estão substituindo os tradicionais termômetros de mercúrio, os medidores de pressão arterial digitais, os oxímetros e os medidores portáteis de glicose são exemplos de dispositivos médicos cada vez mais presentes nos lares.

A Agência Nacional de Vigilância Sanitária – ANVISA - define dispositivos médicos como *“produtos para a saúde, tal como equipamento, aparelho, material, artigo ou sistema de uso ou aplicação médica, odontológica ou laboratorial, destinado à prevenção, diagnóstico, tratamento, reabilitação ou anticoncepção e que não utiliza meio farmacológico, imunológico ou metabólico para realizar sua principal função em seres humanos, podendo entretanto ser auxiliado em suas funções por tais meios”*[2]. Os dispositivos médicos também são conhecidos por termos como correlatos, produtos médicos ou produtos para saúde.

Assim como quaisquer produtos, os dispositivos médicos também estão sujeitos à falhas. Resultados imprecisos, materiais contaminados, presença de corpos estranhos, baixa efetividade e defeitos de fabricação são alguns dos exemplos de falhas relacionadas à esses dispositivos. Um dos casos de falha desse tipo de produto, com grande repercussão mundial, ocorreu com as próteses mamárias da marca francesa *Poly Implant Prothese* (PIP). Após inspeções realizadas na planta de fabricação, a Agência francesa AFSSAPS

descobriu que, desde o ano de 2001, a empresa fabricante estava preenchendo as próteses mamárias com gel silicone diferente daquele aprovado no momento do registro do produto. Em março de 2010, a AFSSAPS determinou a suspensão da comercialização, distribuição, exportação e uso em território francês. Em seguida, o produto foi banido em todo o mercado europeu[3]. No Brasil, a ANVISA suspendeu, por meio da RE 1558/2010, no dia 01 de abril de 2010, "*em todo o território nacional, a comercialização, distribuição, importação e utilização de implantes mamários fabricados pela empresa francesa Poly Implant Prothese (PIP)*"[4].

Tendo em vista que as análises pré-mercado, realizadas para a concessão do registro dos dispositivos médicos, não são capazes de captar todos os problemas que os produtos podem apresentar ao entrar no mercado, o monitoramento desses produtos, após a sua utilização em larga escala, é um dos mecanismos que permite identificar potenciais problemas, principalmente aqueles relacionados à eventos raros e que ocorrem em populações que não foram incluídas nos estudos, como crianças e mulheres grávidas, por exemplo. A ANVISA é o órgão responsável pelo recebimento, processamento e investigação das notificações sobre Eventos Adversos (EA) e Queixas Técnicas (QT) dos dispositivos médicos registrados no Brasil. Essas notificações são lidas e analisadas. Algumas, inclusive, são objeto de investigação.

A ANVISA recebe todos os anos dezenas de milhares de notificações relacionadas à EA e QT relacionadas aos dispositivos médicos. O número de notificações vem crescendo ano após ano, porém, sem o incremento proporcional de recursos humanos necessários para conduzir a leitura, processamento e investigação de tais notificações. Atualmente, todo esse processo é realizado manualmente. Nesse contexto, o desenvolvimento de métodos que permitam a leitura, processamento e geração de informações, a partir desses dados, de maneira automatizada, seria fundamental para a melhor alocação de recursos humanos e para possibilitar uma ação ainda mais tempestiva por parte da Agência, evitando que as falhas relacionadas a algum produto ocasionem maiores transtornos à sociedade.

## 1.1 Definição do Problema

O procedimento para concessão de registro de dispositivos médicos no Brasil segue uma classificação quanto ao risco intrínseco que representam à saúde do paciente, operador ou terceiros. São quatro as classes de risco, onde a Classe I representa a classe de produtos com menor risco e a Classe IV os produtos de maior risco[2].

Os Estados Unidos da América (EUA) utilizam uma sistemática de classificação de risco, para a concessão de registro dos dispositivos médicos, semelhante a classificação de risco adorada pelo Brasil, onde os fabricantes são compelidos a enviar apenas as informa-

ções mínimas necessárias para assegurar a eficácia e segurança dos dispositivos médicos. A diferença é que, ao invés de quatro classes de risco, os EUA adotam três. A Classe I representa os produtos com menor risco e a Classe III os produtos com o maior nível de risco[5]. Como ambos os países possuem a mesma sistemática de registro, é razoável se esperar que enfrentem os mesmos desafios, oportunidades e limitações.

Como uma das consequências da divisão dos produtos em classes de risco, foi observado, nos processos de registro de produtos norte-americano, que muitos dispositivos médicos, classificados nas classes de baixo e moderado risco, são aprovados sem suporte de evidências clínicas[6] e muitos dos dispositivos médicos classificados como de alto risco são aprovados com base em estudos clínicos limitados [7].

Mesmo os dispositivos médicos aprovados com base em estudos considerados robustos, do ponto de vista metodológicos, estão sujeitos a apresentar uma efetividade abaixo da esperada, ou mesmo apresentar eventos adversos que não foram detectados nos estudos apresentados no momento da sua aprovação [7, 8]. Apesar de serem considerados a melhor evidência, existem limitações que são inerentes à esse tipo de estudo, principalmente, quando aplicados à saúde. São limitações relacionadas ao desenho, objetivo da pesquisa, falta de financiamento, monitoramento excessivo, falta de transparência e estruturas inadequadas para a condução do estudo[9].

Diferentemente das consequências provocadas por falhas na maioria dos outros produtos utilizados no dia a dia pelas pessoas, e que não estão diretamente relacionados com à saúde da população, as falhas provocadas pelos dispositivos médicos possuem maior potencial para gerar consequências graves à saúde de seu usuário, podendo levar à lesões permanentes e, em casos mais extremos, levar à óbito[10].

Diante das limitações das avaliações pré-mercado dos dispositivos médicos e das potenciais consequências relacionadas às Queixas Técnicas (QT) e aos Eventos Adversos (EA) que esse tipo de produto pode apresentar, faz-se necessário que haja uma vigilância pós-mercado (VIGIPOS) forte, capaz de detectar tempestivamente essas falhas, agindo para que os danos, na população que está exposta à esse produtos, possam ser minimizados.

O número de notificações recebidas pela ANVISA sobre EA e QT, relacionados aos produtos para saúde, vem crescendo ano após ano, conforme é possível perceber pela Figura 1.1. Conforme exposto no documento “*Tratamento das notificações de evento adverso e queixa técnica relacionadas a produtos para a saúde (equipamentos, materiais e kits para diagnóstico de uso in vitro)*”, disponível no portal eletrônico da Agência<sup>1</sup>, todas as notificações recebidas são lidas e avaliadas, porém, é a gravidade do EA ou o potencial risco da QT que determinam se uma determinada notificação será ou não investigada. A gravidade e a frequência definem a oportunidade da investigação.

---

<sup>1</sup><http://portal.anvisa.gov.br/orientacoes-e-formularios>



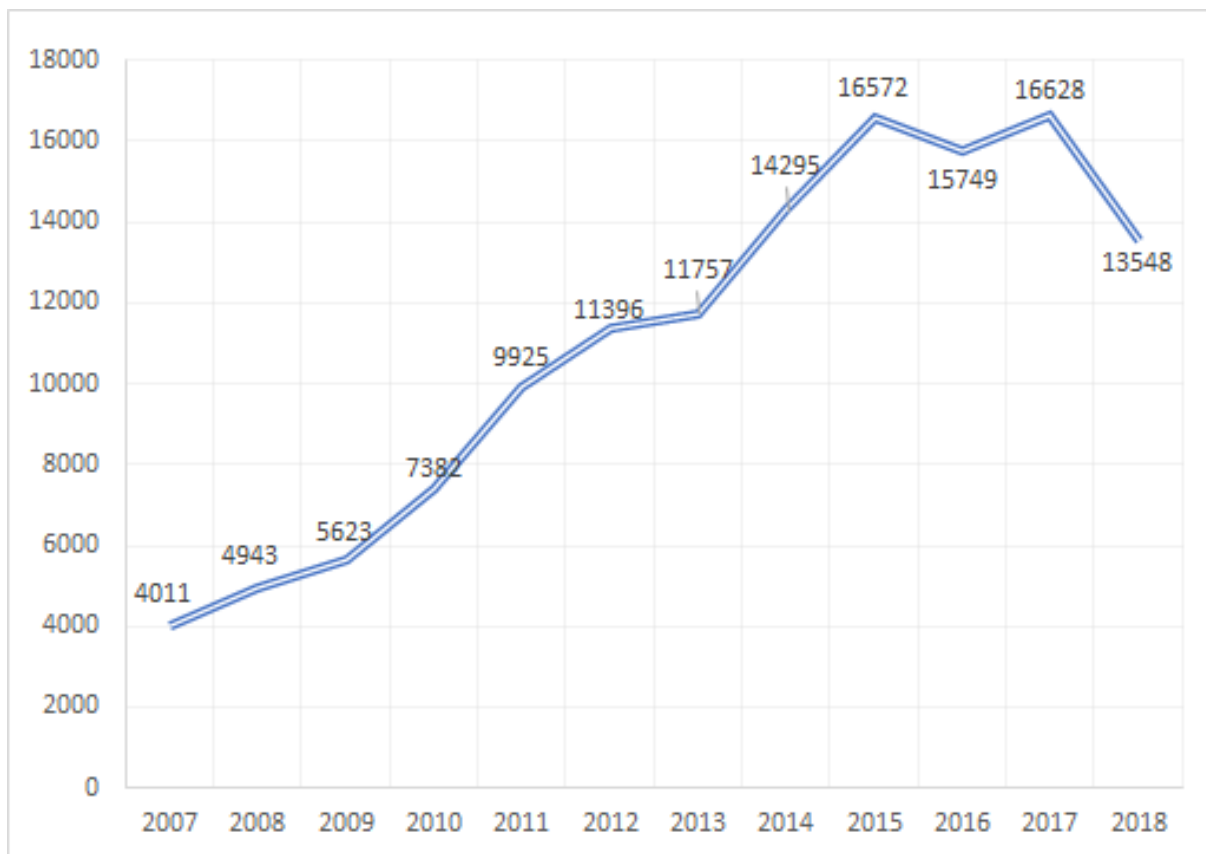


Figura 1.1: Quantidade de notificação ao ano.

O objetivo das investigações é a verificação da relação de causalidade entre o EA ou QT com o produto, com a finalidade de impedir ou minimizar os danos à saúde da população exposta ao produto. Logo, são os resultados das investigações que permitem a ANVISA intervir no mercado para evitar ou minimizar que as disfuncionalidades observadas possam atingir um grupo ainda maior de pessoas.

As notificações relativas à óbito, lesões permanentes ou temporária grave são de investigação obrigatória. Como existem campos específicos, no formulário do sistema NOTIVISA, para que o notificante possa sinalizar se houve óbitos, lesões permanentes ou temporária grave, esses desfechos são automaticamente identificados e separados para investigação.

O maior desafio reside nas notificações classificadas como não graves, ou seja, as notificações que não estão diretamente relacionadas com lesões temporária graves ou óbitos. Nesses casos, a Agência adota como critério, para a realização da investigação, o aumento da frequência associada a esses eventos e à sua capacidade operacional. Porém, essas informações não estão estruturadas, encontram-se em campo de texto livre no formulário, o que não permite uma identificação automática e depende de uma leitura e

processamento manual, pelos técnicos da Agência, para determinar possíveis alterações nas frequências e a necessidade ou não de realização de uma investigação. Como essa atividade concorre com as atividades de investigações relativas às notificações classificadas como graves, relacionadas com óbitos ou lesões, permanentes ou temporárias, e como a capacidade operacional é limitada, é natural que não sejam priorizadas. Assim não é possível captar alterações nas frequências e, conseqüentemente, esses casos não vão para investigação, o que impede a ação tempestiva da Agência para evitar que outras pessoas sejam expostas à esse tipo de EA ou QT classificadas como não grave.

## 1.2 Justificativa do Tema

Importante ressaltar que, apesar de não estarem relacionadas diretamente ao óbito, lesões permanentes ou temporária grave, isso não quer dizer, necessariamente, que essas notificações não possam ocasionar complicações graves à saúde no médio ou longo prazo. Uma notificação que não cause imediatamente algum tipo de dano mais grave, pode ocasionar, indiretamente, ou mesmo de maneira mais lenta, gradual e silenciosa, complicações importantes no estado de saúde de um usuário de um dispositivo médico que apresente um EA ou QT.

Para a captação de alterações nas frequências de EA não graves ou QT, contidas nas notificações, que possibilite o início de investigações e uma atuação tempestiva da ANVISA, seria necessário que todas as notificações fossem lidas, armazenadas e processadas de maneira estruturada e tempestiva, possibilitando a Agência oferecer uma resposta rápida, a fim de se evitar que o EA ou QT notificados venham a se repetir, evitando agravos à saúde em outros grupos de usuários. Atualmente, a Agência não possui um sistema ou processo estruturado para essa captação, de maneira automática, de alterações nas frequências dos EA não graves e das QT, o que acaba prejudicando a identificação de alterações nas frequências e, conseqüentemente, inviabilizando as investigações e a atuação tempestiva da Agência.

A utilização eficiente dos recursos humanos é outro ponto importante. Considerando que um ano possui 252 dias úteis e que um servidor possui direito a um mês de férias, sobrariam 231 dias úteis de trabalho. Considerando que a área possui quatro servidores para analisar as notificações e que pela Figura 1.1, o ano de 2017, que é o último ano com dados completos, teve 16.628 notificações, seriam 18 notificações ao dia para serem analisadas por cada técnicos, por dia, além das outras atribuições como investigações, elaborações de alertas entre outras.

Soma-se a isso, a escassez de mão-de-obra refletida no déficit de servidores da ANVISA. A Agência apresentava um deficit de quase 700 servidores no ano de 2018, somado a um

contingente de mais de 900 servidores em condições de se aposentar <sup>2</sup>. Sem realizar concursos para carreiras de nível superior desde 2013, a Agência perdeu quase 12% da sua força de trabalho entre os anos de 2017 e 2019<sup>3</sup>.

Em um cenário de aumento no número de notificações, conforme dados apresentados pela Figura 1.1, e escassez de mão-de-obra, encontrar uma solução que permita lidar com uma massa cada vez maior de dados é um desafio. A Agência deve buscar a eficiência na execução dos seus processos de trabalho. O desenvolvimento de uma solução para leitura, processamento e análise dessa crescente massa de dados, extraindo entidades-chaves, que permitam captar possíveis variações nas frequências das notificações, tornaria o monitoramento dos EA e QT, classificados como não graves, mais eficiente e menos dependente de recursos humanos, o que permitiria, por sua vez, uma ação tempestiva da Agência, evitando futuros danos à saúde de outras pessoas. A detecção automática de variações nas frequências de EA ou QT, classificadas como não graves, possibilitaria que os recursos humanos, da unidade responsável pela tecnovigilância na Agência, sejam realocados de maneira mais eficiente.

## 1.3 Objetivos

A partir das justificativas apresentadas na Seção anterior, é necessário definir os objetivos do estudo. Os objetivos serão divididos entre objetivo geral, que apresenta a ideia central do trabalho, e os objetivos específicos, que aprofunda o objetivo central e suas peculiaridades.

### 1.3.1 Objetivo Geral

O principal objetivo do presente trabalho é desenvolver uma solução que permita a identificação de alterações nas frequências, em textos livres, das notificações classificadas como não graves, por meio da extração de informações, mais especificamente, da aplicação de técnica de Reconhecimento de Entidades Nomeadas. Essa solução viabilizaria à ANVISA captar, automaticamente, alterações nas frequências das QT relacionadas à determinados dispositivos, permitindo que a Agência possa agir de maneira tempestiva para evitar maiores danos à saúde da população, que seria exposta aos eventos notificados, além de possibilitar uma alocação mais eficiente dos recursos humanos da Agência.

---

<sup>2</sup><https://www.gov.br/anvisa/pt-br/assuntos/noticias-anvisa/2018/barbosa-busca-apoio-para-recompor-quadro-de-servidores>

<sup>3</sup><https://exame.com/brasil/vacinas-contracovid-19-anvisa-tem-deficit-de-servidores-e-corte-de-orcamento/>

### 1.3.2 Objetivos Específicos

- Comparar o desempenho de modelos gerados a partir do português, com modelos gerados a partir da base em português, traduzida para o inglês;
- Testar a utilização de bases pré-treinadas, transfer-learning, para o Reconhecimento de Entidades Nomeadas - REN;
- Avaliar a performance de diferentes atributos *features* nos desempenhos dos diferentes modelos.
- Avaliar a performance de diferentes tipos de representação de segmentos IO, IOB, IOB2, IOE, IOBE e IOBES.

## 1.4 Contribuições Esperadas

A contribuição esperada do presente trabalho é o desenvolvimento de uma técnica de processamento de linguagem natural que permita o reconhecimento de termos chaves, aqui chamada também de Entidades Nomeadas, nos textos livres, contidos nas notificações de EA e QT, classificadas como não graves, dos dispositivos médicos. O emprego dessa técnica para leitura, processamento e identificação de mudanças nos padrões de notificações, deve detectar automaticamente alterações nas frequências de um EA ou QT, permitindo uma alocação mais eficiente dos recursos humanos da Agência, bem como a diminuição no tempo de resposta à esses eventos.

A possibilidade de realocação de recursos humanos da Agência para as atividades mais estratégicas, como a de investigação das notificações, em detrimento da alocação de pessoal com atividades mais operacionais, como a leitura e estruturação das informações contidas em todas as notificações, contribuiria para a promoção e proteção da saúde da população, que é a Missão da ANVISA <sup>4</sup>.

Outra contribuição importante poderá vir dos experimentos conduzidos para verificar a possibilidade da utilização de bases de dados traduzidas por máquina na suplementação do sistema tradicional de vigilância pós-mercado. Como os sistemas tradicionais de vigilância pós-mercado são baseados em notificações espontâneas, acabam sofrendo com as sub-notificações e como o mercado de dispositivos médicos é dominado por produtos importados, ter a possibilidade de utilizar as bases de dados de outros países para suplementar o sistema de vigilância pós-mercado nacional seria um salto importante na área e um reforço considerável para o sistema nacional de tecnovigilância.

---

<sup>4</sup><http://portal.anvisa.gov.br/institucional>

## 1.5 Estrutura do Trabalho

O presente estudo está estruturado da seguinte maneira: o Capítulo 2 apresenta a Revisão de Estado da Arte, contendo o levantamento e análise dos principais estudos na área de extração de informação e reconhecimento de entidade nomeada, aplicados à vigilância pós-mercado. No Capítulo 3, são apresentados os principais conceitos, modelos e as métricas para a avaliação dos modelos que serão utilizados nos experimentos do presente trabalho. No Capítulo 4 são apresentados os experimentos seguindo a estrutura da metodologia para mineração de dados CRISP-DM[8], apresentando a metodologia de pesquisa, o entendimento do negócio e o entendimento dos dados, a preparação dos dados e a modelagem. Nos dois últimos capítulos são apresentados os resultados, análises, bem como as considerações finais contendo os trabalhos futuros.

# Capítulo 2

## Revisão de Estado da Arte

Este capítulo apresenta a revisão de estado da arte da aplicação de aprendizado de máquina para o Reconhecimento de Entidade Nomeadas (NER). No caso da aplicação de NER em tecnovigilância, foram encontrados poucos estudos. Dessa forma, foram incluídos os estudos que utilizaram técnicas de NER em vigilância pós-mercado, no caso, a farmacovigilância que, assim como a tecnovigilância, se propõe ao monitoramento e investigação de eventos adversos e outros problemas, porém, relacionada aos medicamentos. Com o intuito de captar os avanços mais recentes em NER, buscaram-se também os artigos mais atuais, publicados pós 2018, aplicados a qualquer área do conhecimento. Por fim, foram incluídas as revisões sistemáticas e surveys mais recentes em NER. Para a realização desse levantamento foram utilizados alguns elementos da Teoria de Enfoque Meta Analítico Consolidado – TEMAC[11].

O presente levantamento dos artigos para a construção do estados da arte pode ser dividido em quatro partes:

1. Aplicação de NER em tecnovigilância;
2. Aplicação de NER em farmacovigilância;
3. Estudos mais recentes da aplicação de NER em qualquer área do conhecimento;
4. Revisões sistemáticas e surveys mais recentes.

### 2.1 Preparação da Pesquisa

Antes de se iniciar as buscas por artigos na literatura, é necessário que se definam as bases de dados que serão utilizadas como plataforma, bem como o período analisado, os termos e algoritmos de busca. As bases de dados *Web of Science* (WoS), Google Scholar e Scopus, apesar de possuírem suas particularidades, podem ser consideradas bases de

dados estáveis em termos de cobertura de publicações[12]. Dessa maneira, as bases de dados escolhidas, no presente estudo, serão as bases WoS e Scopus. Como os estudos relacionados a aplicação de NER em tecnovigilância e em farmacovigilância são escassos, não foi definida uma restrição de datas. No caso das buscas pelos artigos mais recentes da aplicação de NER, foram definidos os anos de 2020 e 2021 para os artigos e os últimos cinco anos no caso das revisões sistemáticas e surveys.

Definidas as bases de dados e os períodos de pesquisa, o próximo passo é a definição das palavras-chaves, ou descritores, e os respectivos algoritmos de pesquisa, que serão utilizados nas buscas. Por se tratar de um problema relacionado ao Reconhecimento de Entidades Nomeadas (NER), optou-se, para todas as pesquisas, pela utilização do termo em inglês “*named entity recognition*”. As buscas foram realizadas durante o mês de julho de 2021. Abaixo seguem os termos e os algoritmos de pesquisas, definidos para cada uma das quatro partes do levantamento:

1 - Aplicação de NER em tecnovigilância: o termo “*named entity recognition*” foi combinado com os termos “*technovigilance*”, “*medical device*” e “*surveillance*”, uma vez que o termo “*tecnovigilância*” não é consenso na literatura internacional e, muito frequentemente, opta-se pela utilização do termo “*medical device surveillance*”. Como não foi encontrado nenhum estudo, optou-se pela substituição do termo “*named entity recognition*” por uma combinação dos termos “*machine learning*” e “*deep learning*”. A Tabela 2.1 apresenta os algoritmos de pesquisa e os respectivos resultados das buscas nas bases WoS e Scopus.

Tabela 2.1: Quantidade de artigos com NER aplicado à tecnovigilância

<b>Termo de pesquisa</b>	<b>WoS</b>	<b>Scopus</b>
TS = (“ <i>medical device</i> ”)	8.565	64.361
(TS=(“ <i>medical device</i> ”) AND TS=(“ <i>surveillance</i> ”)) OR TS = (“ <i>technovigilance</i> ”)	254	1.915
((TS=(“ <i>medical device</i> ”) AND TS=(“ <i>surveillance</i> ”)) OR TS = (“ <i>technovigilance</i> ”)) AND TS = (“ <i>named entity recognition</i> ”)	0	0
((TS=(“ <i>medical device</i> ”) AND TS=(“ <i>surveillance</i> ”)) OR TS = (“ <i>technovigilance</i> ”)) AND (TS=(“ <i>machine learning</i> ”) OR TS=(“ <i>deep learning</i> ”))	3	10

2 - Aplicação de NER em farmacovigilância: os termos utilizados para a busca de artigos com a aplicação de NER em farmacovigilância foram “*named entity recognition*” e “*pharmacovigilance*”. Os algoritmos e os resultados das buscas estão listados na Tabela 2.2.

A disponibilidade de uma boa quantidade de artigos que aplicam NER à Farmacovigilância, aliada a escassez de artigos contendo a aplicação de NER em Tecnovigilância,

apontam para uma possível lacuna de pesquisa e reforça a importância do presente trabalho.

Tabela 2.2: Quantidade de artigos com NER aplicado à farmacovigilância

<b>Termo de pesquisa</b>	<b>WoS</b>	<b>Scopus</b>
TS = (“ <i>pharmacovigilance</i> ”)	8.077	9.757
TS=(“ <i>pharmacovigilance</i> ”) AND TS = (“ <i>named entity recognition</i> ”)	21	22

3 - Estudos mais recentes da aplicação de NER em qualquer área: no caso da aplicação de NER, em estudos mais recentes, foram combinados os termos “*named entity recognition*” e “*text*”, nos períodos de 2020 e 2021, para a área de “*Computer Science*”, filtrando apenas artigos e revisões. A Tabela 2.3 apresenta os algoritmos de pesquisa e os respectivos resultados das buscas nas bases WoS e Scopus.

Tabela 2.3: Quantidade de artigos sobre NER em 2020 e 2021

<b>Termo de pesquisa</b>	<b>WoS</b>	<b>Scopus</b>
TS=(“ <i>named entity recognition</i> ”) AND TS = (“ <i>text</i> ”)	170	247

4 - Review e survey: as buscas foram realizadas com a combinação dos termos “*named entity recognition*”, “*review*” e “*survey*”. Como o objetivo foi encontrar os artigos mais recentes, considerou-se os anos de 2020 e 2021. O resultado pode ser observado através da Tabela 2.4. Além dos resultados encontrados nessas buscas, foram incluídos algumas “*review*” e “*survey*” que foram referenciados nos artigos mais recentes sobre NER.

Tabela 2.4: Quantidade de revisões e survey sobre NER entre 2020 e 2021

<b>Termo de pesquisa</b>	<b>WoS</b>	<b>Scopus</b>
TI=(“ <i>named entity recognition</i> ”) AND TI = (“ <i>survey</i> ” OR “ <i>review</i> ”)	14	29

Dentre todos os artigos levantados, após a leitura dos títulos e dos resumos, foram selecionados 48 artigos. Desses, 2 tratam da aplicação de NER em Tecnovigilância, 23 sobre a aplicação de NER em Farmacovigilância, 19 estudos mais recentes da aplicação de NER em qualquer área do conhecimento e 4 revisões sistemáticas. Os resultados do levantamento de estado da arte serão apresentados em tópicos, conforme a lista abaixo:

1. Trabalhos Correlatos - aplicação de NER em Tecnovigilância;
2. Aplicações do NER;
3. Categorias do NER;



4. Fatores que podem influenciar nos resultados do NER;
5. Avaliação.

### 2.1.1 Trabalhos Correlatos - aplicação de NER em Tecnovigilância

O Reconhecimento de Entidades Nomeadas é uma das mais importantes tarefas dentro do campo de Extração de Informação. O termo “Entidade Nomeada” foi apresentado, pela primeira vez, em um desafio de extração de informação da conferência MUC-6[13]. O desafio envolvia o reconhecimento de termos relacionados à pessoas, organizações e localizações geográficas em textos livres. O NER é uma ferramenta de pré-processamento de textos que pode ser aplicada em vários desafios como sumarização de textos, tradução, extração de informação, perguntas e respostas[14], bem como aplicada a diferentes domínios.

A Figura 2.1 apresenta o número de publicações anuais, na base Scopus, relacionadas ao NER. Pelo gráfico, que inclui o período entre as anos de 1997 e 2021, é possível perceber um crescente interesse do meio científico sobre o tema. Do total de artigos científicos publicados no período, 59% foram publicados nos últimos cinco anos, sendo que o recorde de publicações anuais foi alcançado no ano de 2020, com 841 publicações, um crescimento de 24% em relação ao ano de 2019 e de 84% em relação ao ano de 2018.

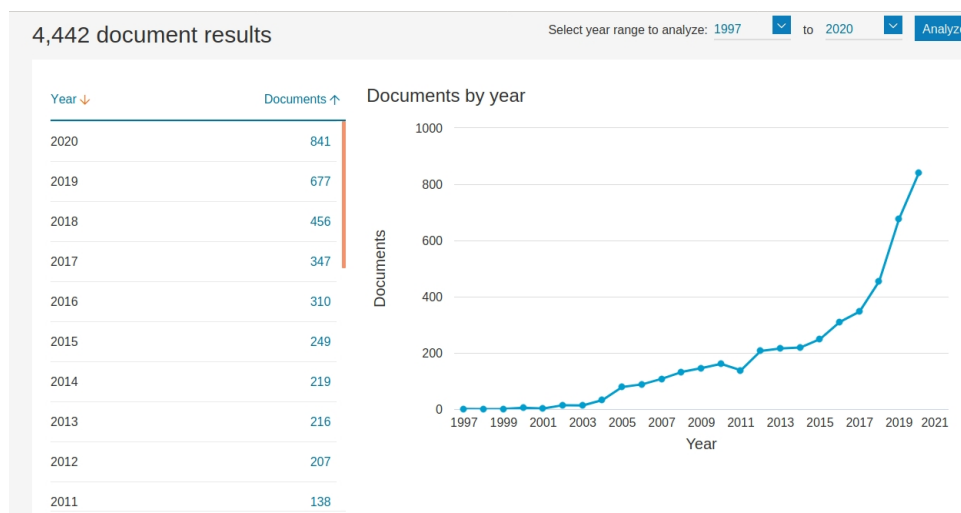


Figura 2.1: Quantidade de publicações ao ano na base *Scopus*.

Apesar do crescente interesse da literatura científica sobre a aplicação de NER em diferentes domínios, a sua aplicação em Tecnovigilância ainda encontra-se incipiente. Dos

estudos encontrados, apenas dois tratavam da aplicação de NER nesse contexto. Ambos serão apresentados a seguir.

Galper et al. (2018)[15] utilizaram processamento de linguagem natural (NLP) para comparar taxas de Eventos Adversos (EA) no uso do *Transcatheter da Válvula Aórtica* (TAVI) e no uso do dispositivo Mitraclip, a partir dos dados contidos, em textos livres, nas bases *Transcatheter Valve Therapy* (TVT) e *The Food and Drug Administration’s Manufacturer and User Facility Device Experience* (MAUDE).

Na fase de pré-processamento, para superar inconsistências e dados faltantes, foram utilizadas técnicas de clusterização e de medidas de similaridade de strings. Para o treinamento do modelo foi utilizado o sistema Boomerang NLP, da empresa Pharma3r LLC, que foi alimentado com bases de treinamento anotadas por um grupo de especialistas. O sistema Boomerang NLP selecionou as *features*, que se dividiam em dois grupos: sintáticas e semânticas. Após o treinamento e classificação das entidades, o desafio foi enquadrar as “prováveis menções” de EA, encontrados nos textos, em labels pré-definidas. Para realizar essa tarefa, o sistema utilizou similaridade entre strings e as relações conhecidas a priori entre EA e dispositivos.

Por fim, os resultados obtidos, através da aplicação da NLP, foram comparados com as frequências de EA e problemas derivados das bases de dados TVT e da base de dados MAUDE. Para exercer a comparação, os autores realizaram uma análise de regressão.

Pelo artigo apresentado, não foi possível identificar quais os modelos e técnicas de *machine learning* que foram empregadas pelo sistema Boomerang NLP para realizar as tarefas de extração dos EA e dos problemas. Apesar de citarem a utilização de técnicas de clusterização, de similaridade de strings, não foi possível captar detalhes sobre as técnicas e sobre os parâmetros utilizados.

Callahan et al. (2019) [16] desenvolveram um método que utilizou uma Rede Neural Recorrente (RNN), do tipo LSTM bidirecional (BiLSTM), para identificar entidades relacionadas ao produto implantado, dor e complicações, nos prontuários médicos eletrônicos, dos pacientes submetidos a implantes de próteses de quadril. Os autores utilizaram um modelo “fracamente supervisionado” (*weakly supervised*), que não necessita de grandes bases anotadas para o seu treinamento.

O objetivo do estudo foi verificar se as evidências extraídas de prontuários médicos eletrônicos poderiam ser utilizadas para tecnovigilância, complementando as evidências obtidas pelo método tradicional, para avaliar a performance dos dispositivos médicos.

Diferentemente do modelo supervisionado, que exige uma base grande de anotações manuais, consumindo tempo e recursos humanos, o modelo fracamente supervisionado necessita de uma pequena base de dados anotada manualmente, que servirá como base de teste para o modelo. No modelo fracamente supervisionado, as palavras contidas

nas anotações clínicas foram transformadas em *tokens* e as entidades foram identificadas, para a base de treino, utilizando um “tagueador” baseado em regras, construído pela combinação de recursos lexicais biomédicos, como o *Unified Medical Language System* (UMLS), e dicionários com curadoria manual. Assim, o modelo gerou uma grande base de treinamento “imperfeitamente anotada”, ou “fracamente anotada”. Por este motivo o modelo é chamado de “fracamente supervisionado”.

Ao treinar um modelo de redes neurais profundas em cima dessa grande base de treino, fracamente anotada, transformam-se as regras dos algoritmos em representação de recursos aprendidas (*learned feature representations*), o que permitiu uma generalização das regras, resultando em um ganho de performance do modelo.

A performance da identificação do modelo foi avaliada em uma base de teste manualmente anotada, que era composta por 500 prontuários anotados manualmente para identificação de menções de implantes, 100 prontuários para identificação de complicações e mais 633 prontuários para identificar menções de dor. O F1 para a identificação de implantes, de complicações e de dor foram de 97.4%, 70.2% e 81.3%, respectivamente. Os autores demonstraram a viabilidade, acurácia e eficiência da utilização de prontuários médicos eletrônicos para a tecnovigilância, através da utilização de um modelo que requer uma quantidade muito menor de dados manualmente anotados para atingir o estado da arte.

O objetivo apresentado pelos supracitados autores é similar ao objetivo do presente estudo, que tem como proposta desenvolver uma solução para a identificação de problemas relacionados aos dispositivos médicos, em textos livres. Uma das diferenças fica por conta da base de dados, já que no presente estudo será utilizada a base de notificação da ANVISA (NOTIVISA) e não os prontuários eletrônicos médicos. Os autores optaram por um modelo fracamente supervisionado, o que demandaria o desenvolvimento de um algoritmo baseado em regras e de dicionários. O modelo utilizado foi o BiLSTM, que é um dos modelos mais utilizados atualmente em NER, conforme será demonstrado a seguir.

Os outros estudos levantados, contendo a aplicação de NER em farmacovigilância, os estudos mais recentes aplicados à qualquer área do conhecimento, bem como as revisões sistemáticas serão apresentados nos tópicos seguintes.

## 2.2 Aplicações do NER

O NER pode ser aplicado nos mais diversos domínios, com os mais diversos propósitos. Dentre as aplicações mais comuns do NER, é possível destacar a extração de informação, perguntas e respostas, sumarização de textos, recuperação de informação, tradução de

textos, mineração de texto, buscas semânticas na web, anotações de vídeos etc[14]. Dentre os artigos levantados, todos os artigos tinham como objetivo a extração de informação.

São muitos os domínios onde o NER vem sendo aplicado. Toledo et al.[17] aplicaram NER para a extração de informação em manuscritos históricos sem a necessidade de um passo intermediário de transcrição. Liu et al.[18] desenvolveram um modelo de computação paralelizada, baseado em *Deep Attention*, e testaram seus resultados em bases de dados genéricas em inglês (CoNLL2003) e chinês (MSRA). Thomas et al.[19] avaliaram seu modelo híbrido em sentenças judiciais. Liao et al.[20] testaram um modelo de redes neurais profundas para encontrar entidades nomeadas relacionadas à tecnologias militares. Além desses domínios, foram encontrados artigos contendo a aplicação de NER na língua indiana[21], neurociência[22] e segurança da informação[23].

Além dos artigos relacionados à Tecnovigilância e a Farmacovigilância, o domínio mais frequente foi a Biomedicina, contendo uma variedade de desafios, com a busca de diferentes tipos de entidades, que incluem entidades de doenças, celulares, clínicas, químicas e medicamentosas. Um total de 7 estudos envolviam a identificação de entidades relacionadas à doenças[24, 25, 26, 27, 28, 29, 30], 4 estudos sobre entidades celulares[31, 32, 27, 30], 3 sobre entidades clínicas[26, 28, 30] e 1 estudo tinha como um dos objetos a identificação de entidades medicamentosas[29].

No caso dos artigos que aplicam NER em Farmacovigilância, a grande maioria dos estudos tem como uma das entidades nomeadas as Reações Adversas(RA)[33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46]. Além das RA, outras entidades encontradas estavam relacionadas à sintomas[47, 34, 36, 37, 48, 38, 45, 46], medicamentos[47, 33, 34, 37, 38, 49, 41, 42, 43, 44, 45, 46] e doenças[36]. No caso da Tecnovigilância, Callahan et al.[16] tinham como objetivo extrair entidades relacionadas ao produto, dor e complicações relacionadas aos procedimentos médicos.

Pelos levantamentos realizados, é possível perceber que o NER é aplicado sobre diversos domínios, com a finalidade de se classificar diferentes entidades. As entidades nomeadas dependem dos objetivos de cada estudo, não havendo uma padronização.

## 2.3 Métodos de NER

Sun et al.[50] e Xu et al.[24] dividem os métodos de Reconhecimento de Entidades Nomeadas em três categorias: Baseados em Regras (“Rule-based”), Baseado em Dicionários (“Dictionary-based”) e Baseados em Aprendizado de Máquina (“Machine Learning-based”). Kanimozhi et al.[51] acrescentam uma quarta categoria, a Híbrida.

O método Baseado em Regras, como o próprio nome sugere, baseia-se em regras feitas a mão por especialistas. O resultado desse método é obtido através da busca,

pelos algoritmos, nos textos livres, retornando os elementos que se encaixam nas regras estabelecidas. Uma limitação importante desse método é que ele é muito dependente do domínio, não sendo possível generalizá-lo, já que as regras são específicas[51] para cada contexto. Por uma série de limitações desse método, que inclui a não generalização, ele tem sido preterido em detrimento do método baseado em Aprendizado de Máquina[14]. Nenhum dos artigos levantados adotou esse método.

O método baseado em Dicionários tenta cruzar os elementos de um texto com os elementos de um dicionário[51]. Esse método é amplamente utilizado nas anotações e indexação nos domínios biomédicos. No entanto, dada a grande variedade de termos existente, é muito difícil que um único dicionário consiga obter bons resultados. É importante ressaltar que tanto o método baseado em regras quanto o método baseado em dicionários são dependentes de um especialista seja para construir as regras ou mesmo os dicionários[14]. Nenhum dos artigos levantados utilizou o método puramente baseado em dicionários.

O método baseado em Aprendizado de Máquina utiliza diferentes tipos de algoritmos para a extração de entidades, de textos livres, sem a necessidade de um dicionário ou de regras preestabelecidas[51]. Esses métodos aprendem pelas *features* dos textos e são menos dependentes do domínio, ou seja, possuem uma capacidade de generalização maior. Porém, esse método costuma depender de um número grande de textos anotados para o seu treinamento[48]. Ele pode ser dividido em três tipos: supervisionado, semi-supervisionado e não supervisionado. Dentre os 48 artigos selecionados para a leitura, 36 artigos aplicavam algum método para a extração de entidades nomeadas em domínios específicos. Os outros 12 artigos tratavam de revisões sistemáticas, surveys ou artigos que pesquisavam aspectos relacionados às *features*, relato de conferências ou que apresentavam novas bases de dados pré-treinadas. Dentre os 36 artigos que aplicavam algum método de NER, 30 utilizavam o método baseado em Aprendizado de Máquina.

Desses 30 artigos, nenhum artigo optou pelo método não-supervisionado, 1 artigo experimentou a utilização de dois métodos, um supervisionado e outro não supervisionado[27] e 2 outros artigos utilizaram métodos semi-supervisionados. Para tornar o processo de anotação mais fácil, Morlane et al.[47] empregaram um *bootstrapping* no classificador. Começaram anotando 100 mensagens e classificando as outras 100. Após uma validação manual, as 100 mensagens classificadas entravam na base de treino. Esse processo continuou até que a qualidade da base anotada para treino fosse considerada aceitável. Chen et al.[38] empregaram um método semi-supervisionado denominado *tri-training*, que se baseia em um mecanismo de votação, por três classificadores diferentes, para selecionar os elementos, da base não anotada, que iriam compor a base de treinamento.

Os outros 27 estudos são do tipo supervisionado, ou seja, necessitam de uma massa

representativa de dados, anotados manualmente, para o seu treinamento. Com relação aos modelos utilizados por esses estudos, houve o predomínio das Redes Neurais, presentes em 24 estudos. Pela opção da utilização de um software privado, não foi possível captar o modelo utilizado em um dos estudos[15] e outros dois estudos utilizavam os modelos de Aprendizado de Máquina clássicos como Árvore de Decisão, Regressão Logística, *Support Vector Machine*(SVM)[48] e *Conditional Random Fields*(CRF)[37].

Considerando apenas os estudos que empregaram Redes Neurais, o tipo mais frequente são variações das Redes Neurais Recorrentes (RNN), como a *Long short-term memory* (LSTM), principalmente, a sua variação Bidirecional (BiLSTM), e a *Gated Recurrent Unit* (GRU). Ao todo, 21 estudos utilizaram a BiLSTM, seja sozinha ou seja em combinação com Redes Neurais Convolucionais (CNN), com mecanismos de atenção ou com uma camada de CRF. Dentre os estudos que não utilizaram o BiLSTM, um estudo desenvolveu uma arquitetura com o GRU e uma camada de CRF[25], outro apresentou uma mistura de Atenção + CNN[32] e um terceiro estudo usou o tipo *Extreme Learning Machine* (ELM), que é um algoritmo de aprendizado eficiente, que tem chamado atenção, principalmente, pela sua velocidade[29].

Com relação aos estudos que utilizavam a arquitetura LSTM, todos optaram pela sua variação, o bidirecional LSTM, a BiLSTM. Por se tratar de textos, o BiLSTM permite o aprendizado de informações nas duas direções para a classificação de uma entidade. A rede BiLSTM pode ser empregada sozinha ou em conjunto com outras redes e modelos. Apenas 3 estudos aplicaram o BiLSTM isoladamente. O BiLSTM foi empregado por Toledo et al.[17] para a extração de informação de documentos antigos, escritos a mão; por Xia et al.[33] para a identificação de reações adversas em textos de mídias sociais; e por Zhang et al.[40] que utilizou o BiLSTM como linha de base na comparação com um modelo de aprendizado multi-tarefa, que utilizava seq2seq + atenção na identificação de reações adversas.

A arquitetura mais frequente nos estudos levantados foi a combinação do BiLSTM com o CRF. Nesse caso, o BiLSTM possui o papel de capturar o contexto no qual a palavra está inserida. O CRF é utilizado para prever a sequência das *labels*[24, 38]. Assim, os modelos se complementam, sendo capazes de, em conjunto, obter um desempenho melhor. Ao todo, 9 artigos utilizaram a arquitetura BiLSTM+CRF pura[36, 39, 49, 26, 21, 27, 28, 30, 22, 42, 41, 43, 44], outros 4 artigos incluíram o mecanismo de Atenção nessa arquitetura - BiLSTM + CRF + Atenção[18, 24, 31, 20, 45, 46] e outros 2 estudos que combinaram o BiLSTM + CRF + CNN[19, 23].

No caso dos estudos que empregaram o método baseado em Aprendizado de Máquina, a maioria é do tipo supervisionado e utilizava modelos de Redes Neurais, sendo a arquitetura BiLSTM + CRF a mais utilizada, seja isoladamente ou em conjunto com mecanismos de

Atenção ou com arquiteturas CNN.

O método híbrido consiste na combinação de mais de um dos métodos apresentados anteriormente. Acaba por utilizar os pontos fortes de cada um dos métodos e tende a obter melhores resultados do que as técnicas individuais[14]. Foram encontrados 6 estudos que optaram pelo método híbrido, sendo 2 deles no domínio da Farmacovigilância[34, 35], 1 para Biomedicina[24], 1 para textos jurídicos[19], 1 para neurociência[22] e 1 em Tecnovigilância[16]. Callahan et al.[16] utilizaram um método baseado em regras e dicionário para a anotação das entidades. Segundo os autores, esses métodos conseguiram atingir marcadores altos de *Precision* e *Recall*.

Como o objetivo do presente estudo é a aplicação de NER para a extração de informação nas bases de notificação da ANVISA, optou-se pela utilização do método, do tipo, do modelo e da arquitetura mais utilizadas na literatura. Assim, conforme os supracitados achados, optou-se pela utilização do método baseado em Aprendizado de Máquina, do tipo supervisionado, e um modelo de Redes Neurais Recorrentes com a arquitetura BiLSTM + CRF.

## 2.4 Fatores que influenciam os resultados em NER

Os resultados da aplicação do NER podem ser influenciados por uma série de fatores. Além dos desafios impostos por um determinado domínio, ou mesmo a dificuldade de se encontrar qual o melhor modelo a ser utilizado, existem outros fatores que influenciam no resultado final. Alguns dos exemplos desses fatores são a língua, disponibilidade de bases pré-treinadas, disponibilidade de bases anotadas, a escolhas das *features* e a representação de segmento.

### 2.4.1 Língua

A língua é um fator que impacta no resultado final. Línguas que são pouco exploradas, que possuem poucos recursos disponíveis ou que possuem características de construção complexas, tendem a oferecer desafios maiores para a aplicação do NER[14]. Um exemplo desses recursos é a disponibilidade de bases pré-treinadas para a utilização em Transferência de Conhecimento (*transfer-learning*), que vem apresentando bons resultados na aplicação de NER nos mais diversos domínios. A língua Inglesa é a mais utilizada no mundo e, por consequência, a mais estudada também, o que acaba oferecendo aos pesquisados um número grande e variado de bases pré-treinadas. No caso da estrutura linguística, por exemplo, a língua Inglesa e muitas línguas europeias utilizam letras maiúsculas para denominar um nome próprio de pessoas, de um país ou de uma cidade, característica que

outras línguas como a chinesa, árabe e hindi não possuem[14]. Esse tipo de estrutura linguística, obviamente, acaba por favorecer o reconhecimento de entidades nomeadas desse tipo.

Dos 36 artigos analisados, 23 artigos tinham como objeto a identificação de entidades nomeadas em Inglês, 3 artigos aplicavam NER em bases de dados em Inglês e Chinês[18, 25, 23], 1 na língua inglesa e espanhola[49], 4 artigos empregaram NER em domínios na língua chinesa[38, 25, 26, 20], 1 artigo na língua espanhola[17], 3 da língua francesa[47, 34, 35] e 1 da língua indiana[21]. Nenhum dos artigos levantados aplicava o NER na língua portuguesa.

### 2.4.2 *Features*

As *features* são características que são utilizadas como *inputs* nos modelos de NER para a identificação das entidades desejadas. No passado, vários métodos se baseavam em *features* manuais e concentravam boa parte dos seus esforços na sua seleção[33], que costuma ser uma das maiores etapas do processo de desenvolvimento do melhor modelo[37]. Existe uma grande variedade de *features* que pode ser utilizadas. Alguns exemplos são as *features* semânticas, sintáticas, morfológicas, ortográficas, lexicais e vetores numéricos que representam caracteres (*character embedding*), vetores que representam palavras (*word embedding*) e que representam um texto (*sentence embedding*).

Os três artigos que utilizaram modelos clássicos de aprendizado de máquina foram os que utilizaram as *features* semânticas, sintáticas, morfológicas, ortográficas ou lexicais[47, 48, 37]. Os artigos que aplicaram redes neurais optaram pela utilização por *features* de vetores numéricos que representavam caracteres, palavras ou trechos do texto. Apenas dois artigos utilizaram *features* diferentes. Toledo et al.[17], que tinham como objetivo a extração de informações de textos históricos escritos a mão, utilizaram como *features* as imagens e os labels das palavras anteriores e Xu et al.[24] utilizaram como *features* as *Pos taggings* e as representações numéricas das palavras. Dentre as representações numéricas, a mais utilizada foi a representação de palavras. Ao todo, 25 estudos utilizaram essa representação como uma das *features* dos seus modelos. 14 estudos utilizaram representações de caracteres e 5 estudos empregaram vetores representativos de trechos dos textos. Conforme observado no levantamento realizado, e pelo modelo se tratar de uma rede neural, pretende-se utilizar *features* numéricas de palavras, caracteres ou trechos dos textos.



### 2.4.3 Representação de Segmento

Existem entidades nomeadas que são formadas por mais de uma palavra, ou por palavras compostas. A forma de representação desse tipo de entidade é um desafio e costuma impactar no resultado final. Ao longo dos anos, foram desenvolvidas diferentes maneiras de se representar as entidades em um texto, o que ficou conhecido na literatura como representação de segmento *segment representation*. Nessas representações, cada uma das palavras, ou *tokens*, são classificadas com um *label* que se inicia com uma letra específica. As representações vão desde as mais simples como a IO chegando até as mais complexas, como a BIOES[14]. A Tabela 2.5 apresenta algumas representações de segmento encontradas na literatura.

Tabela 2.5: Algumas Representações de Segmento disponíveis na literatura

<b>Representação de Segmento</b>	<b>Significado Labels</b>
IO	<i>Inside e Outside</i>
IOB	<i>Inside, Outside e Begin</i>
IOE	<i>Inside, Outside e End</i>
IOBE	<i>Inside, Outside, Begin e End</i>
BILOU	<i>Begin, Inside, Last e Outside e Unit</i>
BIOES	<i>Begin, Inside, Outside e End e Single</i>

Konkol et al.[52] realizaram uma série de experimentos para comparar as diferentes representações de segmentos. Nos experimentos eles aplicaram dois modelos de aprendizado de máquina, em quatro línguas diferentes, e encontraram que as representações de segmento, assim como a língua e a abordagem utilizada, influenciam no resultado final. Concluíram, ainda, que a representação que obteve o melhor desempenho foi a IOE (*Inside, Outside e End*). Ratnov et al.[53] confrontaram duas representações de segmento, a BILOU e a BIO (IOB), sendo esta última uma das mais utilizadas na literatura. Ao final, a BILOU superou significativamente a performance da BIO.

Dentre os artigos levantados, 20 deles utilizaram algum tipo de representação de segmento. 10 artigos utilizaram a representação BIO[24, 36, 38, 40, 19, 25, 26, 20, 22, 23, 42, 41, 43, 44, 45, 46], 2 artigos optaram pela representação BIOES[18, 31], 1 pela BMEWOV[49], 1 pela IOU[29] e 1 pela representação BIO2[30]. O presente estudo tem como um de seus objetivos avaliar a performance dos diferentes tipos de representações de segmento na identificação das entidades de interesse.

## 2.5 Avaliação

Existem diferentes maneiras para se mensurar a performance do NER na literatura. Em resumo, o que essas métricas tentam captar é a capacidade dos modelos em identificar

corretamente as entidades. Para realizar essa checagem, costuma-se confrontar as predições realizadas pelos modelos com as predições realizadas por humanos, normalmente, através de uma base de dados anotada. Métricas como *Precision*, *Recall* e *F1-score* são as mais utilizadas para realizar essa comparação[14]

Essas métricas costumam ser aplicadas a uma base de dados anotada. Essa base deve ter uma estrutura pré-definida e utilizar padrões aceitáveis para poder ser utilizada para treinamento, validação e teste dos modelos desenvolvidos. A construção desse corpus tende a exigir muito esforço e tempo e, normalmente, é executada por um especialista. Os corpus padrão-ouro (*Gold Standard Corpus* - GSC) é uma coleção de textos anotados manualmente. Um GSC de alta qualidade requer a participação de vários especialistas, que realizam anotações independentes, onde as divergências são equacionadas para assegurar qualidade. Porém, fatores como a ambiguidade dos dados e nível de conhecimento dos especialistas podem influenciar no resultado final[54].

Os artigos utilizaram GSC previamente construídos por outros pesquisadores e disponíveis na literatura ou criaram seus próprios corpus anotados. A grande maioria dos estudos levantados utilizou GSC próprios para realizar seus experimentos. Callahan et al.[16] desenvolveram um modelo “fracamente supervisionado”, onde utilizaram um algoritmo baseado em regras e em dicionário para realizar anotações no corpus de treinamento, enquanto que a base de teste foi composta por 500 textos anotados manualmente. Chen et al.[34] desenvolveram um guia para a anotação dos textos, que foram realizadas de maneira independente por dois especialistas em farmacovigilância. O tamanho do GSC foi definido em 45 textos por conceito. Molane-Hondere et al.[47] escolheram anotar manualmente 1.200 textos selecionados aleatoriamente. Afirmaram, ainda, que essa quantidade é superior a quantidade de textos anotados em estudos correlatos.

Weber et al.[27] apresentam uma lista de GSC biomédicos, que tratam de diferentes tipos de entidades, como entidades químicas, celulares, genes, espécies e doenças. Dentre os artigos que utilizaram GSC disponíveis na literatura, Sun et al.[32] utilizaram a base GENIA, que contém entidades celulares e a base ACE, essa última composta por entidades de localização, veículos, armas etc. Mais recentemente, Wei et al.[42], Dandala et al.[41], Yang et al.[43] e christopoulou et al.[44] utilizaram o corpus n2c2 MIMIC-III, composto por 505 textos anotados, contendo informações sobre medicamentos.

A maioria dos autores utiliza a métrica de F1-score, aplicada a uma base fixa de teste e treino, para realizar seus experimentos e comparar diferentes modelos. Porém, Chen et al.[38] compararam três modelos por meio da utilização de um método de reamostragem por bootstrap, por 20 vezes, com diferentes bases de treinamento. Assim, os três modelos foram aplicados nas 20 bases de treino, sendo extraídos os valores das médias e dos desvios padrão das métricas de *Precision*, *Recall* e *F1-score*. Os autores utilizaram os dados das

médias e dos desvios padrão do *F1-score* para avaliar os modelos. Wunnava et al.[39] realizaram seus experimentos na base MADE 1.0, que possui entidades relacionadas à medicamentos e eventos adversos. Para confrontar seus modelos, utilizaram um teste t pareado para testar a significância estatística das diferenças dos F1-score obtidos para cada uma das entidades, bem como para testar o F1-score geral.

O presente estudo se propõe a desenvolver um modelo que será testado em um corpus próprio, formado por uma amostra aleatória da base de dados de notificações da ANVISA. A base será anotada manualmente e o modelo será treinado e testado, a partir de 20 amostras formadas por reamostragem por bootstrap, e será aplicado um teste t pareado para verificar se as diferenças obtidas na métrica *Macro F1-score* possuem significância estatística, já que essa é a métrica de avaliação mais utilizada na literatura.

# Capítulo 3

## Fundamentação Teórica

Nesse capítulo serão apresentados os principais conceitos relacionados à extração de informação e ao Reconhecimento de Entidade Nomeadas (NER). Primeiramente, será apresentada a metodologia CRISP-DM, utilizada para auxiliar na condução de projetos de mineração de dados. Em seguida, como a intenção é utilizar o NER, serão apresentados ainda os conceitos básicos de Processamento de Linguagem Natural (NLP), extração de informação, NER e os modelo *Conditional Random Fields* e BiLSTM + CRF, bem como os principais indicadores de avaliação dos modelos de NER.

### 3.1 CRISP-DM

O processo de Mineração de Dados é complexo e dependente da utilização de boas ferramentas e da habilidade dos analistas que o conduzem. Sem um processo bem definido não é possível saber quais as etapas e, o mais importante, qual o seu resultado final. Nesse contexto que surge o CRISP-DM (*Cross Industry Standard Process for Data Mining*), que tem como objetivo solucionar parte desse problema ao propor uma estrutura de processo para os projetos de mineração de dados, ajudando a traduzir problemas reais em desafios de mineração de dados. Seu objetivo é tornar esses projetos mais eficientes, confiáveis, gerenciáveis e replicáveis[8].

O CRISP-DM é uma metodologia descrita em termos de um modelo de processo hierárquico, propondo a divisão do projeto de mineração de dados em seis fases: entendimento do negócio, entendimento dos dados, preparação dos dados, modelagem, avaliação e implementação[8]. A Figura 3.1 ilustra as diferentes fases do CRISP-DM.

1. **Entendimento do Negócio:** corresponde a primeira fase do projeto. Nessa fase, o analista procura conhecer o problema e os objetivos sob a perspectiva do negócio com vistas a delimitar o problema de mineração de dados e desenhar um plano de ação preliminar para atingir os objetivos desejados.

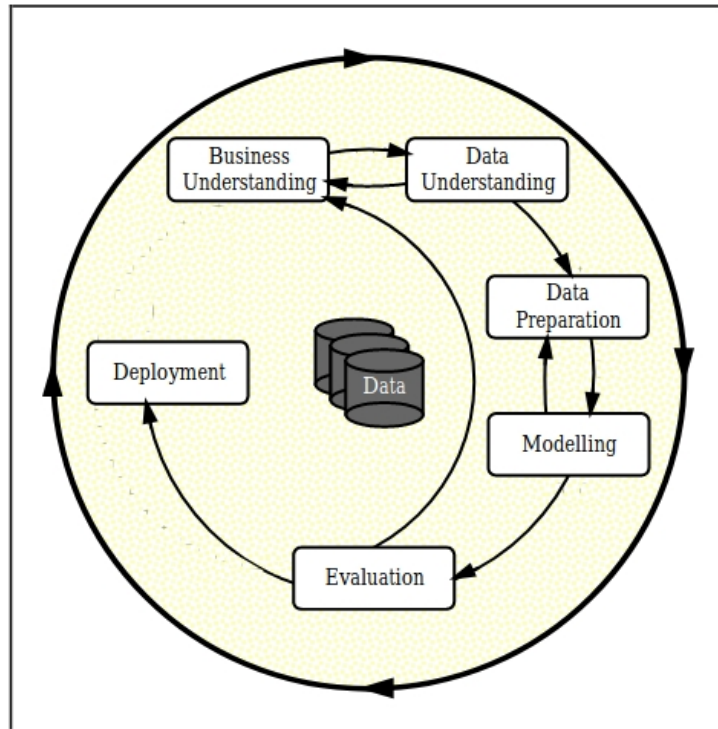


Figura 3.1: Fases do CRISP-DM (Fonte: [8]).

2. **Entendimento dos dados:** começa com uma coleta inicial dos dados. O objetivo é obter uma familiarização com os dados, identificar problemas de qualidade de dados, obter *insights* e a realizar a primeira busca por hipóteses que podem servir de base para o estudo. Existe uma ligação importante entre as fases de entendimento do negócio e de entendimento dos dados, já que para a formulação do problema de mineração de dados e do plano de ação é necessário algum entendimento dos dados.
3. **Preparação dos dados:** cobre todas as atividades necessárias para a formação do conjunto de dados que será utilizado no modelo, a partir dos dados iniciais brutos. A preparação de dados pode acontecer em diversos momentos do processo e pode ser repetido quantas vezes for necessário. As atividades que estão englobadas nessa fase são a tabulação, catalogação, gravação, limpeza e transformação dos dados para as ferramentas de modelagem, bem como a seleção e construção de novos atributos.
4. **Modelagem:** normalmente, existe mais de uma técnica possível para um mesmo problema de mineração de dados. Técnicas que exigem formatos específicos de dados. Nessa fase, várias técnicas de modelagem são selecionadas e aplicadas aos dados, prevendo, ainda, os ajustes de parâmetros necessários para a busca de resultados ótimos. Essa fase está relacionada com a fase de preparação de dados, na medida

em que problemas relacionados ao conjunto de dados ou mesmo a necessidade de novos dados são, frequentemente, percebidas no momento da modelagem.

5. **Avaliação:** nessa fase espera-se que o melhor ou os melhores modelos já tenham sido selecionados sob a perspectiva dos resultados obtidos na fase de modelagem. Antes de se partir para a implementação do modelo, é necessário que haja uma avaliação do modelo e dos passos utilizados para se chegar aos resultados encontrados, com vistas a ter certeza que os objetivos pretendidos inicialmente foram adequadamente atingidos. O objetivo principal dessa fase é descobrir se teve algum fator importante do ponto de vista do negócio que não foi adequadamente considerado. Espera-se que ao final dessa etapa a decisão de se utilizar ou não os resultados encontrados na fase de modelagem seja tomada.
6. **Implementação:** o conhecimento adquirido ao longo de todo o processo de mineração de dados precisa ser organizado e apresentado para que o cliente possa utilizá-lo. O processo de implementação depende do contexto em que se está inserido. Assim, o resultado final pode variar desde a simples apresentação de um relatório final até uma implementação de um processo de mineração de dados permanente.

## 3.2 Processamento de Linguagem Natural - NLP

O NLP surgiu nos anos 50 como uma interseção entre a inteligência artificial e a linguística. Inicialmente, o NLP surgiu como um campo separado da recuperação de informação (*information retrieval text*), que se baseia em técnicas estatísticas escaláveis para indexar e procurar um grande volume de texto de maneira eficiente. Com o tempo, o NLP foi convergindo com uma série de outros campos, incluindo o próprio campo de recuperação de informação[55].

Essa convergência vem possibilitando o desenvolvimento de uma série de tecnologias. Dentre as tecnologias mais difundidas, é possível destacar a extração de informação (em inglês, *information extraction*), reconhecimento de entidade nomeada (em inglês, *named entity recognition*), rastreamento de tópicos (em inglês, *topic tracking*), sumarização (em inglês, *summarization*), categorização (em inglês, *categorization*), clusterização (em inglês, *clustering*), ligação de conceito (em inglês, *concept linkage*), visualização de informação (*information visualization*), entre outras[56]. O sítio eletrônico NLP-progress<sup>1</sup> apresenta os principais desafios de NLP, com os respectivos estados da arte em cada um deles.

Como o objetivo do presente estudo é a extração de informações-chaves nos textos livres referentes às notificações de EA e QT, relativas aos dispositivos médicos no Brasil,

---

<sup>1</sup><http://nlpprogress.com/>

dentre as tecnologias de NLP supracitadas, duas serão melhor exploradas: extração de informação e reconhecimento de entidade nomeadas.

### 3.2.1 Extração de Informação

Normalmente, esse é o ponto de partida para que os computadores possam analisar textos desestruturados. Uma das tarefas da extração de informação é identificar frases chaves e as relações que existem dentro de um texto, através de um processo conhecido como correspondência de padrões. A extração de informação é uma tecnologia muito útil para lidar com um grande volume de textos[56].

As principais tarefas relacionadas à extração de informação são: (1) identificação de erros gramaticais ou de digitação (*Spelling/grammatical error identification and recovery*), (2) desambiguação do sentido das palavras (*Word sense disambiguation*), (3) negação e identificação de incerteza (*Negation and uncertainty identification*), (4) extração de relacionamento (*Relationship extraction*), (5) inferências temporais (*Temporal inferences*) e (6) reconhecimento de entidade nomeada (*named entity recognition*)[55].

### 3.2.2 Reconhecimento de Entidades Nomeadas - NER

O NER procura identificar frases ou palavras (“entidades”) em um texto, com o objetivo de categorizá-los. Exemplos de tipos de entidades são: pessoas, locais, organizações, medicamentos, doenças etc. Uma tarefa muito comum, relacionada ao NER, é o mapeamento de entidades nomeadas para conceitos em um vocabulário. Porém, algumas situações fazem com que o NER seja um desafio, são eles[55]:

- **Variação na ordem das palavras:** as frases podem ser escritas na voz passiva ou na voz ativa, o que altera a ordem das palavras nas frases com os mesmos sentidos;
- **Palavras derivadas:** os sufixo e prefixos das palavras são alguns dos exemplos;
- **Inflexão:** o tempo verbal, o plural, as diferentes conjugações verbais, superlativo etc;
- **Sinônimos:** a utilização de palavras diferentes com o mesmo significado atrapalha na identificação das entidades;
- **Homografias:** palavras que apresentam a mesma grafia, mas com pronuncia e significados diferentes;

O NER é a tarefa mais fundamental no ramo de mineração de textos biomédicos. A identificação de entidades relacionadas a identificação das proteínas, espécies, doenças e

componentes químicos são alguns dos exemplos de aplicação dessa tecnologia. Nos dias atuais, as ferramentas que obtêm os melhores resultados em NER são as que utilizam *features* específicas, na tentativa de capturar características específicas, relacionadas a alguma entidade que se queira identificar. A descoberta desses atributos (*features*) é mais uma arte do que uma ciência, necessitando de um processo extensivo de tentativa e erro. Muitas das ferramentas mais eficientes de NER exigem uma lista, ou dicionário, o que é difícil e custoso para se manter[57].

Por outro lado, os métodos utilizados para o reconhecimento de entidades nomeadas, em um dado conjunto de texto, com base em *features* pré-definidas, são bastante homogêneos. Um dos métodos mais utilizados, considerado como o método padrão, para a definição de uma linha de base, é o *Conditional Random Fields* (CRF)[57].

### **Atributos ou *features***

As *features* é o insumo mais importantes nas tarefas de NER. A maioria das tarefas de NER se baseiam em *features* específicas para a captura de classes específicas das entidades nomeadas. Exemplos de *features* pode ser a própria palavra, a palavra anterior, a palavra posterior, a *POS-tagging* dessas palavras, o sufixo da palavra, o prefixo da palavra, entre outras. A depender do contexto, uma *feature* faz mais sentido que outra. Por exemplo, a utilização de sufixo em tarefas de NER relacionadas ao reconhecimento de nomes de moléculas de medicamentos tem potencial para ser uma *feature* importante, já que os nomes de muitos medicamentos terminam com as mesmas sequências de caracteres. Exemplo: omeprazol, pantoprazol e esomeprazol.

A engenharia de *features* é muito importante para a obtenção de bons resultados e está mais para arte do que propriamente para ciência. Encontrar as *features* que ajudam a obter um melhor resultado é um processo de tentativa e erro, sendo muito custoso para o processo de montagem do modelo[57]. Dentro da tarefa de NER, uma das mais importantes, e mais pacificada das *features*, é a *POS-tagging*.

### ***Word Embedding***

A partir da aplicação de redes neurais para as tarefas de NER, outras *features* foram sendo testadas. Conforme visto no Capítulo 2, as mais utilizadas na literatura são os vetores numéricos que representam caracteres (*character embedding*), vetores que representam palavras (*word embedding*) e que representam um texto (*sentence embedding*). Dentre essas, a mais utilizada são os vetores que representam palavras (*word embedding*).



### *Part of speech tagging - POS-tagging*

POS tagging é um problema muito conhecido nas tarefas de PLN e consiste em assinalar uma tag morfossintática (substantivo, verbo, adjetivo, pronome, pontuação etc.) para cada uma das palavras de um texto. Como muitas palavras podem apresentar ambiguidade, O *POS-tagging* pode ser um meio para ajudar a amenizar essa questão, uma vez que seu objetivo é classificar cada uma das palavras levando em consideração as regras gramaticais e, por consequência, o contexto no qual a palavra se insere[58].

Existe uma série de classificadores de *POS-tagging*, uma série de corpus para treinamento e uma série de sistemas de *POS-tagging*. Desde o primeiro sistema de tagueamento, TAGGIT[59], lançado em 1971, muitos avanços tem sido alcançados na busca de classificadores com alto nível de acurácia e eficiência[58]. O *POS-tagging* é um insumo importante na utilização de outras tarefas de PLN, tais como *parsing* e tradução[60]. Nas tarefas relacionadas a REN, o *POS-tagging* é utilizado como uma das *features* que auxiliam na identificação da classe da entidade.

### **3.3 Avaliação do modelo proposto**

Em estudo recente, Goyal et al.(2018) apontou que existem diferentes métricas utilizadas na literatura para avaliar o desempenho dos modelos. Encontraram que as avaliações possuem o foco em medir a capacidade dos modelos em encontrar corretamente as entidades. As métricas mais utilizadas pelos estudo foram *Precision*, *Recall* e *F1-measure*[14]. Esses indicadores são calculados a partir das métricas extraídas da matriz de confusão, ilustrada na Tabela 3.1.

Tabela 3.1: Matriz de confusão

	<b>Previsão (+)</b>	<b>Previsão (-)</b>
<b>Observação (+)</b>	Verdadeiro Positivo (VP)	Falso Negativo (FN)
<b>Observação (-)</b>	Falso Positivo (FP)	Verdadeiro Negativo (VN)

No caso do Reconhecimento de Entidades Nomeadas, por exemplo, o Verdadeiro Positivo (VP) ocorre quando uma palavra é classificada como pertencendo a determinada classe de Entidade, quando a palavra realmente pertence a classe assinalada. O Falso Negativo (FN) acontece quando se afirma que a palavra não pertence a determinada classe de Entidade quando na verdade ela pertence a classe. Analogamente, o Falso Positivo (FP) ocorre quando se afirma que a palavra pertence a determinada classe quando ela pertence a outra classe. Por último, o Verdadeiro Negativo (VN) ocorre quando se afirma que a palavra não pertence a determinada classe quando ela realmente não pertence aquela classe.

### 3.3.1 *Precision*

O *Precision* tem como objetivo captar a capacidade do modelo em prever, dentre todas as previsões positivas, qual foi o índice de acerto. A fórmula é a seguinte:

$$Precision = \frac{VP}{(VP + FP)} \quad (3.1)$$

### 3.3.2 *Recall*

O *Recall* tem como objetivo captar a capacidade do modelo em prever, dentre todas as observações efetivamente positiva, qual foi o índice de acerto. A fórmula é a seguinte:

$$Recall = \frac{VP}{(VP + FN)} \quad (3.2)$$

### 3.3.3 *F1-Score*

O *F1-Score* é calculado a partir de uma média harmônica entre o *Precision* e o *Recall*. Em situações onde há um grande desbalanceamento das classes, situação muito habitual na prática, o índice acurácia pode ser alto com um *Recall* baixo, o que significa um modelo com bons resultados, mas com pouca utilidade. Nesses casos, o F1-Score entrega uma métrica que faz mais sentido ao considerar no seu cálculo o *Precision* e o *Recall*.

$$F1 - Score = \frac{2 * (Recall * Precision)}{(Recall + Precision)} \quad (3.3)$$

A maioria dos estudos de NER envolve mais de uma entidade nomeada. Dessa maneira, é necessário que se mensure a performance do modelo no geral, bem como para cada uma entidade. Assim, existem duas variações de métricas de *F1-score* que podem realizar essa tarefa: *Macro-averaged F1-score* e *Micro-averaged F1-score*. A *Macro-averaged F1-score* é a média de todas as *F1-score* das entidades do corpus, enquanto que a *Micro-averaged F1-score* é calculada após o agrupamento de todas as entidades de um corpus e das métricas de *Precision*, *Recall* e *F1-score*. A *Micro-averaged F1-score* é muito afetada pelos parâmetros da maior classe[14], assim, no presente trabalho optou-se pela utilização da *Macro-averaged F1-score* para a definir o desempenho dos modelos.

A estimativa desses parâmetros é um passo importante não apenas para prever a "acurácia" do modelo, mas também para a seleção do modelo. Mais importante do que saber o valor do parâmetro estimado é conhecer o viés e a variância. Dois métodos de reamostragem comumente utilizados para essa finalidade são o *Cross-validation* e o *Bootstrap* [61]. Em seguida, serão apresentados esses métodos que serão utilizados ao longo do estudo.

## Bases de treino e teste

Utilizar a mesma base de dados para realizar o treinamento e o teste do modelo é um equívoco metodológico, uma vez que o modelo estaria sendo treinado para prever classes que ele conhece, prejudicando a generalização de seu resultado. Essa situação é conhecida como *Overfitting*.

Uma prática muito utilizada para evitar o *Overfitting* é a divisão da base em duas partes, uma destinada para o treinamento do modelo e outra destinada para o teste do modelo. A proporção entre treino e teste pode variar. As proporções mais utilizadas variam de 70%-30% a 90%-10% para treino e teste, respectivamente. Normalmente, uma proporção maior é separada para o treinamento do modelo.

### *Cross-validation*

Quando existe a necessidade de se ajustar os hiperparâmetros do modelo, a utilização de uma base de treino e uma de teste não evita *Overfitting*. Isso acontece, pois os hiperparâmetros podem ser ajustados até que o modelo obtenha a melhor performance na base de teste, perdendo a capacidade de generalização.

Nesse caso, o recomendável é a divisão da base de dados em três partes: treino, validação e teste. Dessa forma, o modelo pode ser treinado na base de treino, testado na base de teste, com posterior avaliação pela base de validação. Após o ajuste dos hiperparâmetros, o modelo pode obter uma avaliação final na base de teste. No entanto, a divisão da base de dados em três partes faz com que a base de dados para treinamento se reduza e com que o resultado seja altamente dependente das escolhas aleatórias das bases de treino e validação.

Uma solução para esse tipo de problema é a utilização do *Cross-Validation*(CV). A solução mais básica é o k-fold CV. A base é dividida em duas partes, uma de treino e outra de teste. No k-fold CV, a base de treino é subdividida em k partes. O modelo é treinado com base em k - 1 partes e o resultado é validado com a parte remanescente dos dados. As medidas de performance é computada a partir da média dos valores computados nas k rodadas. O resultado final, após o ajuste dos hiperparâmetros, é computado na base de teste, separada no início do processo. Atualmente, o valor mais utilizado na literatura é de 10-fold, onde  $k = 10$ .

### *Bootstrap*

A partir de uma amostra de tamanho n, a amostragem por *Bootstrap* prevê a criação de outras amostras de mesmo tamanho n, a partir da seleção aleatória, com distribuição uniforme e reposição, de elementos da amostra original. Assim, o método de *bootstrap*

permite a obtenção de um bom estimador do parâmetro real da população a partir da reamostragem de apenas uma amostra.

No caso da reamostragem de 20 amostras com reposição de tamanho  $n$ , obtidas a partir da amostra original, é possível obter uma estimativa de um parâmetro de interesse a partir da média dos valores dos parâmetros obtidos em cada uma das 20 amostras.

### 3.4 *Conditional Random Fields - CRF*

O CRF foi proposto, primeiramente, por Lafferty et al. (2001) [62]. CRF generaliza a regressão logística para dados sequencias assim como *Hidden Markov Model - HMM* generaliza o *Nayve Bayes*[55]. CRF possui a habilidade de prever múltiplas variáveis que dependem uma da outra. Possui uma aplicação diversa como a classificação de regiões em imagens, classificação de gene no DNA e aplicações em PLN, além do *POS-tagging* e REN[63].

A ideia é prever um vetor de resultados  $y = \{y_0, y_1, \dots, y_T\}$  de variáveis aleatórias, dado um vetor de *features*  $x$ . No caso da tarefa de *POS-tagging*, cada uma das variáveis  $y_s$  é a *POS-tagging* que se deseja prever da palavra na posição  $s$ , enquanto que os *inputs*  $x$  são divididos em vetores de *features*  $\{x_0, x_1, \dots, x_T\}$ . Cada um dos inputs  $x_s$  contém várias *features*, sobre a palavra na posição  $s$ . *Features* como a identidade da palavra, características ortográficas, sufixos, prefixos, informações semânticas, entre outras[63].

De acordo com (Sutton and McCallum, 2012), um CRF de cadeia linear é uma distribuição condicional que segue a seguinte equação:

$$p(y|x) = \frac{1}{Z(x)} \prod_{t=1}^T \exp \left\{ \sum_{k=1}^K \Theta_k f_k(y_t, y_{t-1}, x_t) \right\} \quad (3.4)$$

Onde  $Z(x)$  é uma função de normalização dada pela equação:

$$Z(x) = \sum_y \prod_{t=1}^T \exp \left\{ \sum_{k=1}^K \Theta_k f_k(y_t, y_{t-1}, x_t) \right\} \quad (3.5)$$

Habibi et al.(2017)[57] compararam o CRF puro com LSTM-CRF e com algoritmos específicos de NER que melhor desenvolveram em tarefas específicas. Uma das conclusões que os autores chegaram é que o método CRF puro possui uma performance superior aos algoritmos específicos. Os autores encontraram duas explicações para esse fenômeno. A primeira é que muitos dos algoritmos são desenhados para tarefas muito específicas, para uma base de dados específica, fazendo com que os mesmos resultados não sejam alcançados quando esses mesmos modelos são extrapolados para outras bases de dados, em outros contextos. Outra explicação é que no experimento que conduziram, utilizaram

o CRF com *word embeddings*, o que possibilitou ao modelo reconhecer similaridades semânticas entre as palavras.

### 3.5 Long-Short Term Memory

LSTM significa Long Short Term Memory, em uma tradução livre para o português, significaria Memória de Curto e Longo prazos, e é um tipo de Redes Neural Recorrentes (RNN). A RNN é uma rede amplamente utilizada para solucionar problemas sequenciais como, por exemplo, séries temporais ou mesmo textos, que são formados por um conjunto de palavras que, ordenadamente, geram sentido ao texto.

A RNN clássica consegue guardar representações recentes de entradas na forma de ativações, ou memória de curto-prazo, porém, não consegue guardar a memória de longo prazo. Enfrentam uma dificuldade para carregar informações de uma sequência longa. Essas redes sofrem com a dissipação do gradiente (*vanish gradient*). Os gradientes são utilizados para atualizar os pesos da rede. reduzem consideravelmente, ao longo da propagação, chegando ao ponto de deixarem de contribuir para a atualização dos pesos.

Para solucionar esse problema, foram introduzidas novas *features* na RNN clássica para formar o LSTM, conhecidos como "portões". Um portão de entrada foi criado para proteger a memória armazenada. O portão de saída foi criado para proteger outras unidades de informações irrelevantes. Existe ainda a memória da célula (*memory cell*), que recebe informações de estados passados e dos portões de entrada e saída[64]. A Figura 3.2<sup>2</sup> ilustra uma representação dos principais portões da arquitetura LSTM.

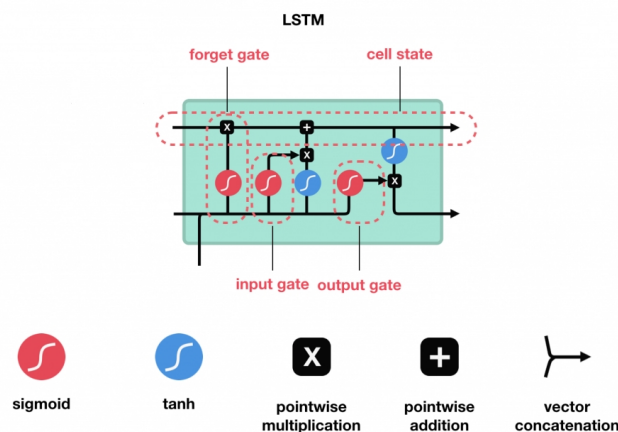


Figura 3.2: Portões que regulam a passagem de informações.

<sup>2</sup>ilustração retirada de <https://www.deeplearningbook.com.br/arquitetura-de-redes-neurais-gated-recurrent-unit-gru/>

# Capítulo 4

## Experimentos

Este Capítulo apresenta a metodologia de pesquisa utilizada, os experimentos realizados e os resultados alcançados. A realização dos experimentos, assim como os tópicos do Capítulo, foram divididos pelos seis estágios da metodologia CRISP-DM[8].

### 4.1 Metodologia de Pesquisa

O primeiro estágio corresponde ao **entendimento do negócio**, onde se busca a compreensão dos objetivos e pré-requisitos da perspectiva do negócio, convertendo em um problema de mineração de dados e um plano inicial para se atingir tais objetivos. Nesse estágio serão apresentadas informações obtidas por meio de entrevistas com dois técnicos da área e por meio de um levantamento de literatura, com a finalidade de definir o objetivo do trabalho.

O segundo estágio é o de **entendimento dos dados**, que corresponde a coleta inicial dos dados, visando a familiarização e entendimento dos mesmos, buscando encontrar problemas de qualidade, bem como coletar os primeiros *insights*. Nesse estágio será realizada uma exploração inicial da base de dados NOTIVISA com o objetivo de explorar a maneira como os dados são extraídos e quais suas principais características. Será realizada uma estatística descritiva para tirar as primeiras informações, entender como a base está dividida, quais as classes de produtos com mais notificações, quais os tipos de notificações mais frequentes, o número de notificações por gravidade e quais os produtos mais notificados.

Na fase de **pré-processamento** são definidos os dados que serão utilizados na fase de modelagem, a partir dos dados apresentados na etapa de entendimento dos dados. Nessa etapa serão apresentados os dados que serão selecionados para compor a base de dados e para a criação do corpus; os recursos utilizados na extração, armazenamento e

processamento dos dados; as classes e entidades de interesse; e os passos utilizados para a criação do corpus.

Na etapa de **modelagem** são definidos não apenas o modelo, como também são definidos os atributos que serão utilizados e os tipos de experimentos que serão realizados com o intuito de encontrar o melhor modelo. Essa fase contempla a limpeza dos dados, tratamento de dados faltantes, chegando até a definição dos atributos.

No momento da avaliação, o(s) modelo(s) selecionados são avaliados de forma mais completa, garantindo que ele(s) atinja(m) os objetivos do negócio. No último estágio, o de desenvolvimento, o conhecimento ganho deverá ser organizado e apresentado de uma maneira onde o cliente possa utilizá-lo.

## 4.2 Entendimento do negócio

Para se obter um melhor entendimento do negócio, foram realizadas, no ano de 2018, entrevistas com dois técnicos que trabalham na área de tecnovigilância da ANVISA. Além das entrevistas, foi realizado um levantamento de literatura para complementar as informações obtidas por meio das entrevistas.

A maior parte dos produtos para saúde são considerados bens credenciais, ou seja, os consumidores desses bens não são capazes, por si só, de atestarem a sua qualidade, necessitando de um agente credenciador para a realização desse ateste[65]. A Anvisa, além de ser responsável pela concessão do registro desses produtos, também é a Agência responsável pela coleta, processamento e manutenção das notificações sobre Eventos Adversos (EA) e Queixas Técnicas (QT) relativas aos medicamentos e dispositivos médicos.

Enquanto processos regulatórios rigorosos são aplicados para a concessão de registro de um medicamento, através da exigência de ensaios clínicos randomizados (ECR) que sejam capazes de avaliar a eficácia e segurança dos medicamentos, o mesmo não ocorre para a aprovação de registro dos dispositivo médico[66]. Na área da saúde, os ECR são considerados como o padrão ouro para o levantamento de evidências sobre a eficácia e segurança dos produtos e serviços da saúde, perdendo apenas para as revisões sistemáticas com metanálise[67]. Apesar de serem considerados a melhor evidência, não estão livres de limitações e barreiras. As principais dizem respeito ao desenho inadequado dos estudos, compreensão inadequada do objetivo da pesquisa, falta de financiamento, monitoramento excessivo, falta de transparência e estruturas inadequadas para a condução do estudo[9].

Mais especificamente para os dispositivos médicos, existem outras limitações para a condução dos ECR para além daquelas já apontadas anteriormente. Bernard et al.[68] conduziu uma revisão sistemática da literatura com o objetivo de levantar problemas relacionados à condução de ECR para dispositivos médicos. Seus achados apontaram para

a existência de, pelo menos, cinco barreiras para a condução de ECR para dispositivos médicos, são elas: tempo de avaliação, aceitabilidade, cegamento do estudo, escolha do grupo comparador e a curva de aprendizado.

Soma-se a essas limitações para o levantamento de evidências sobre a eficácia e segurança dos dispositivos médicos, no momento do registro, a habilidade limitada dos profissionais de saúde, individualmente, na detecção do aumento de problemas relacionados aos dispositivos médicos, após o início de sua comercialização. Dois fatores contribuem para essa limitação: o pequeno número de procedimentos que realizam e a dificuldade na implementação de um programa de monitoramento rigoroso dos pacientes após a utilização dos dispositivos. A detecção de eventos adversos mais raros, bem como o aumento da frequência de eventos adversos esperados requerem processos sistematizados de vigilância em larga escala[69].

Nesse contexto surge a Tecnovigilância, Vigpós, ou vigilância pós-mercado de dispositivos médicos, que é definida pela ANVISA como *“o sistema de vigilância de eventos adversos e queixas técnicas de produtos para a saúde na fase de pós-comercialização, com vistas a recomendar a adoção de medidas que garantam a proteção e a promoção da saúde da população”*[70]. Seu principal objetivo é verificar a segurança dos dispositivos médicos pós-comercialização, com o foco na ocorrência de eventos adversos, queixas técnicas, falsificação de produtos e situações de séria ameaça à saúde.

A Resolução da diretoria colegiada – RDC nº 67, de 21 de dezembro de 2009 dispõe *“sobre as normas de tecnovigilância aplicáveis aos detentores de registro de produtos para saúde no Brasil”*. Dentre as principais regras, é possível destacar:

1. Tornar obrigatória a notificação de eventos adversos e queixas técnicas pelo detentor de registro de dispositivos médicos;
2. Estabelecer os prazos máximos para notificações;
3. Obrigatoriedade do detentor de registro de dispositivo médico de designar, por documento escrito, um profissional responsável pela área de tecnovigilância da empresa;
4. Necessidade de estruturação e implantação de um sistema de tecnovigilância pelas empresas detentoras de registro de dispositivos médicos.

Os eventos adversos são definidos pela RDC nº 67, de 21 de dezembro de 2009 como *“qualquer efeito não desejado, em humanos, decorrente do uso de produtos sob vigilância sanitária”*. A RDC divide, ainda, os eventos adversos em graves e não-graves. Os eventos adversos graves são aqueles que podem levar o paciente à óbito, requerer intervenção médica para prevenir o comprometimento permanente, exigir ou prolongar a hospitalização do paciente ou levar ao risco fetal, morte fetal ou a uma anomalia congênita. As



queixas técnicas, por sua vez, são definidas como *qualquer notificação de suspeita de alteração/irregularidade de um produto/empresa relacionada a aspectos técnicos ou legais, e que poderá ou não causar dano à saúde individual e coletiva*[70].

As notificações são coletadas e armazenadas pelo Sistema Nacional de Notificação de Eventos Adversos e Queixas Técnicas (NOTIVISA). O sistema NOTIVISA foi implantado em 2007. As informações são coletadas via web e armazenadas na plataforma Oracle. O NOTIVISA permite acesso em tempo real pelos habilitados para receberem as informações, no caso, os componentes do Sistema Nacional de Vigilância Sanitária, que é formado pela ANVISA e pelos entes responsáveis pela Vigilância Sanitária nos estados e municípios[71]. Além da coleta e armazenamento das notificações obrigatórias das empresas detentoras de registro de dispositivos médicos, o NOTIVISA também recebe notificações voluntárias de hospitais, profissionais de saúde, cidadãos e quaisquer pessoas ou organizações que queiram notificar um evento adverso ou queixa técnica.

As notificações recebidas pelo NOTIVISA são lidas e avaliadas de maneira manual. Não existindo nenhum processo automatizado que auxilie no processamento das notificações recebidas e armazenadas no banco de dados da Agência. A decisão sobre a investigação ou não de uma notificações é definida pela gravidade do evento adverso ou pelo potencial de risco da queixa técnica. As notificações de óbito, lesões permanente ou temporária grave são de investigação obrigatória. As outras notificações permanecem no banco de dados e serão investigadas quando a análise de tendência definir que sejam investigadas. Nesses casos, é adotado o critério de aumento da frequência associado à capacidade operacional da unidade técnica da ANVISA para investigação. O principal objetivo das investigações é verificar a existência de causalidade entre o incidente notificado com a finalidade de minimizar ou impedir a probabilidade de ocorrência dos mesmos eventos nos usuários desse produto. Essas ações podem resultar em modificações no registro do produção, suspensão da fabricação ou mesmo o cancelamento do registro.

Dessa forma, considerando a competência da ANVISA para exercer a atividade de Tecnovigilância; a vasta e crescente massa de dados recebidas e armazenadas, todos os dias, pelo sistema NOTIVISA; a necessidade de urgência nas ações relacionadas às notificações de eventos adversos e queixas técnicas; e a limitação da investigação pela capacidade operacional da unidade técnica, o objetivo do presente projeto é a construção de um modelo de aprendizado de máquina que seja capaz de extrair informações das notificações, que seja capaz de identificar eventos adversos e queixas técnicas associadas a cada um dos produtos e que possibilitem a identificação de variações nas frequências.

## 4.3 Entendimento dos dados

O entendimento dos dados é a fase em que se obtém acesso à base de dados e se inicia uma análise preliminar dos dados. Os objetivos iniciais são verificar a qualidade dos dados e buscar hipóteses que possam ser testadas durante o estudo. Em conjunto com a fase de entendimento do negócio, a fase de entendimento dos dados é o momento em que os projetos de Mineração de Dados estão sendo definidos e conceitualizados. Quando os objetivos e os recursos disponíveis são estabelecidos e a viabilidade do projeto é avaliada. Caso ocorram mudanças estruturais e significativas, o projeto deve ser reiniciado.

O NOTIVISA (Notificação de Eventos Adversos e Queixas Técnicas) é o sistema informatizado, desenvolvido pela ANVISA em 2007, para a coleta e armazenamento das notificações relacionadas aos Eventos Adversos(EA) e às Queixas Técnicas(QT) sobre produtos e serviços sob vigilância sanitária, incluindo os dispositivos médicos (artigo médico-hospitalar e equipamento médico-hospitalar), medicamentos, vacinas, kit reagente para diagnóstico in vitro, assistência à saúde, cosméticos, sangue ou componente, saneantes e agrotóxicos. Cada um dos produtos ou serviços possuem seus próprios formulários, com estruturas específicas para a coleta das notificações. No caso do presente estudo, os formulários que serão objeto de estudo serão quatro:

1. Formulário 1 - Evento Adverso associado ao uso de Artigo Médico-Hospitalar (Anexo I);
2. Formulário 2 - Queixa Técnica de Artigo Médico-Hospitalar (Anexo II);
3. Formulário 3 - Evento Adverso associado ao uso de Equipamento Médico-Hospitalar (Anexo III);
4. Formulário 4 - Queixa Técnica de Equipamento Médico-Hospitalar (Anexo IV);

Os quatro formulários eletrônicos somam 195 campos. Além dos campos necessários para a caracterização do EA ou QT, os outros campos se referem à identificação do produto, da empresa fabricante ou importadora, do notificador onde ocorreu a QT ou EA, bem como informações adicionais para melhor entendimento acerca do contexto em que o EA ou QT estão inseridos. A quantidade de campos existente em cada formulário estão listadas na Tabela 4.1.

### 4.3.1 Estrutura dos formulários

Todos os quatro formulários possuem algumas seções em comum, dentre elas, a identificação do motivo da notificação (EA ou QT), a identificação do produto, identificação da empresa responsável pelo produto, outras informações e a seção de pendências. Uma

Tabela 4.1: Quantidade de campos por formulário

<b>Formulário</b>	<b>Nº de Campos</b>
Formulário 1	56
Formulário 2	39
Formulário 3	59
Formulário 4	41

seção que é exclusiva dos formulários de notificação de EA é a seção de identificação do paciente. Nessa seção do formulário são coletados os dados dos pacientes que sofreram os Eventos Adversos e é dividida em duas partes, sendo uma referente às informações do paciente, ou usuário, e outra referente à evolução do caso, ou evento adverso. Os campos referentes às informações do paciente contém, além dos campos para a coleta de informações básicas do paciente (nome, sexo, raça e nascimento etc), dados do prontuário e da internação médica, quando for o caso. A parte que coleta os dados necessários para o acompanhamento da evolução do caso possui três campos que são utilizados para determinar se um EA é classificado como grave ou não grave. São campos que coletam dados sobre a ocorrência de óbito ou lesão.

A seção de identificação do produto e da empresa produtora, ou importadora, além dos dados cadastrais dos produtos e das empresas, ainda coletam dados acerca da fabricação do produto, como lote, data de fabricação, data de validade entre outros. A seção de outros motivos coleta outros dados sobre a situação que envolveu a Queixa Técnica ou Evento Adverso como a existência de nota fiscal, amostras ou rótulos do produto, o local da aquisição do produto entre outras. A seção de pendência lista todas as necessidades de correções no preenchimento dos formulários, caso existam.

Na seção Motivo é onde são detalhadas as QT ou EA. No caso das QT, é necessário que o notificante enquadre em um dos cinco motivos existentes: (1) Produto com suspeita de desvio de qualidade, (2) Produto com suspeita de estar sem registro, (3) Suspeita de empresa sem autorização de funcionamento (AFE), (4) Suspeita de produto falsificado ou (5) Suspeita de outras práticas irregulares. No caso dos EA, há a necessidade de classificação dos mesmos conforme a terminologia de reações adversas da Organização Mundial de Saúde (*World Health Organization – Adverse Reaction Terminology*).

Ainda na seção Motivo, há um campo aberto para a descrição detalhada da QT ou do EA, incluindo a descrição dos motivos da notificação. Nesse campo, o notificante relata, com maiores detalhes, como ocorreu a QT ou o EA. Os dados provenientes desse campo servirão de insumos na presente pesquisa. Como as notificações ocorrem em textos livres, não estruturados, haverá a necessidade da identificação das entidades referentes as QT ou EA para, em seguida, contabilizar alterações nas suas frequências e permitir que as QT ou EA classificados como não graves sejam objeto de investigação.

### 4.3.2 Estatísticas das notificações

A base de dados que será utilizada no presente estudo compreende os dados coletados entre 1 de janeiro de 2007 e 31 de outubro de 2018. Os dados foram extraídos do sistema e disponibilizados em formato *Comma-separated values* (CSV). Os mesmos estão estruturados em formato de tabela, onde as linhas representam os registros e as colunas representam os campos dos formulários. Ao todo, são 131.829 registros com 106 colunas (campos), uma vez que nem todos os campos foram extraídos do sistema.

Ao realizar uma análise prévia dos dados, as 131.829 notificações sobre EA ou QT, ocorridas no período, são referentes à 10.470 produtos, de 2.339 empresas. As tabelas abaixo apresentam algumas estatísticas extraídas da base de dados:

Tabela 4.2: Tipo de notificação

<b>Tipo</b>	<b>%</b>
Queixas Técnicas	85,53%
Eventos Adversos	14,47%

Tabela 4.3: Percentual de notificações por classe de produto

<b>Classe de produto</b>	<b>%</b>
Artigos médico-hospitalar	92%
Equipamentos médico-hospitalar	7%
Kit diagnóstico	1%

Tabela 4.4: Notificações por gravidade

<b>Notificações</b>	<b>%</b>
Graves	8,80%
Não Graves	91,20%

A partir dos dados apresentados nas Tabelas 4.2, 4.3 e 4.4, há um domínio de queixas técnicas, de notificações sobre artigos médicos-hospitalares e a grande maioria das notificações são classificadas como não graves, que serão objeto do presente estudo. A Tabela 4.5 mostra que foram 11.601 notificações classificadas como graves. A maioria diz respeito às lesões temporárias graves e a minoria provocaram óbitos.

A Tabela 4.6 apresenta os produtos com mais notificações no período. Dentre esses, os Dispositivos Médicos Implantáveis (DMI) são os que apresentam o maior risco para a saúde dos pacientes, uma vez que são inseridos dentro do corpo. Dentre os DMIs, o Catéter é o DMI com o maior número de notificações, seguido pelos implantes dentários. A Tabela 4.7 apresenta as 10 classes de notificantes que mais realizaram notificações. É

Tabela 4.5: Notificações por tipo de gravidade

<b>Tipo</b>	<b>Notificações</b>
Óbito	1.540
Lesões Permanentes	2.742
Lesões Temporárias Graves	7.319

Tabela 4.6: Os 10 produtos com mais notificações

<b>Produto</b>	<b>Notificações</b>
Equipos	10.204
Seringas Descartáveis	10.435
Catéter	8.601
Implantes Dentários	8.329
Luvras Descartáveis	6.797
Sondas	4.878
Compressas	4.183
Luvras Cirúrgicas	3.752
Agulhas Descartáveis	3.012
Torneiras	2.055

possível verificar que a metade das notificações foram realizadas pelos hospitais sentinelas e que mais de 92% das notificações foram realizadas por hospitais ou pelas empresas.

No caso dos dados referentes as descrições dos motivos das notificações, coletados nos campos de textos livres, e que serão objeto de estudo dessa pesquisa, os dados não são uniformes. Existe uma grande variação no tamanho dos textos das notificações. O menor texto possui apenas 4 caracteres, enquanto que os maiores possuem 40.000 caracteres. A média do número de caracteres é de 194 caracteres, enquanto a mediana é de 112 caracteres. A Figura 4.1 mostra o histograma com a distribuição de frequências das quantidades de caracteres das notificações. Para fins de se obter uma melhor visualização, o histograma foi construído com as notificações compostas por até 1.116 caracteres.

No caso das notificações selecionadas para investigação, aproximadamente, 60% foram concluídas. Os status do restante das investigações estão descrito na Tabela 4.8. As investigações que são concluídas são divididas em seis categorias, tanto para os EA quanto para as QT. As seis categorias são: inconclusivo, descartado, confirmado, provável, improvável e possível[71]. Porém, para fins de tecnovigilância, cabem apenas quatro dessas possibilidades: **Confirmado**, **Provável**, **Inconclusivo** e **Descartado**.

1. **Confirmado**: No caso de EA, os casos confirmados acontecem quando é confirmada uma relação causal entre o evento adverso e o uso do produto. No caso das QT, ocorre quando a QT notificada é confirmada;

Tabela 4.7: As 10 classes de notificantes que mais notificaram

Notificantes	Notificações
Hospitais Sentinelas	65.026
Hospitais	31.234
Empresas	25.179
Estabelecimentos de saúde	1.691
Serviços de hemoterapia	848
Núcleos de segurança do paciente	738
Universidades/centros de pesquisa	332
Vigilâncias sanitárias municipais	308
Operadoras de Planos de Saúde	290
Vigilâncias sanitárias estaduais	238

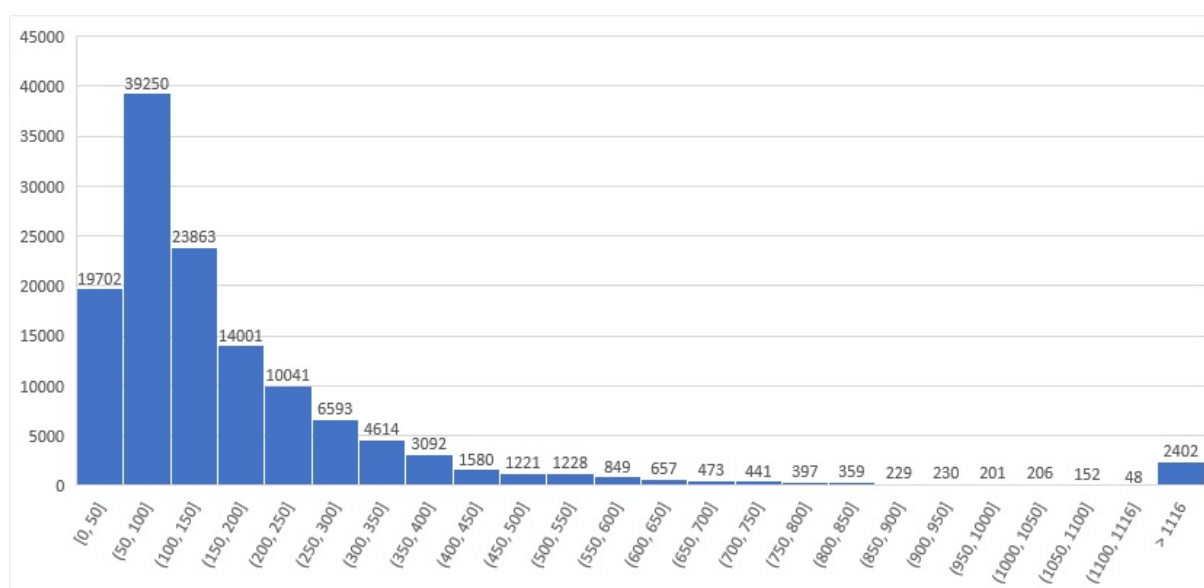


Figura 4.1: Histograma quantidade de caracteres das notificações de QT e EA.

2. **Provável:** No caso de EA, quando não foi confirmada, mas existem evidências sobre a relação causal entre o uso do produto e o evento notificado. Para as QT, quando não foi confirmada, mas existem evidências de irregularidade ou desvio de qualidade;
3. **Inconclusivo** - quando as informações levantadas na investigação não são suficientes ou são contraditórias, não sendo possível a confirmação da relação causal entre o uso do produto e o EA, ou não sendo possível a confirmação da QT;
4. **Descartado** - No caso de EA, quando for confirmada a inexistência de relação causal entre o uso do produto e o evento adverso. No caso da QT, quando não for confirmada a QT.

Tabela 4.8: Status das notificações em investigação

Status	%
Em agrupamento	15,59
Em análise	14,79
Em investigação	8,21
Outros	1,41

Obs: 40% das notificações selecionadas para investigação

## 4.4 Preparação dos dados

No presente estudo, serão testados dois modelos de aprendizado de máquina supervisionado, um modelo base, o *Conditional Random Fields* (CRF), que será confrontado com uma Rede Neural Recorrente, com uma arquitetura BiLSTM + CRF, que tem apresentado bons resultados na literatura. Serão utilizados dois corpus, um com as notificações originais em português e outro com as notificações traduzidas para o inglês. O objetivo é verificar o impacto da features e das bases pré-treinadas no resultado final. Serão ainda realizados experimentos com a utilização de features pré-treinadas e com a utilização de diferentes tipos de representação de segmento.

### 4.4.1 Dados selecionados

Dentre os campos disponíveis nos formulários de notificação, os campos selecionados para utilização no presente trabalho serão os campos de “*código da notificação*”, “*tipo de notificação*”, “*descrição da ocorrência*” e o “*nome técnico do produto*”. Todos os outros campos não serão considerados na fase de modelagem.

O campo “*código da notificação*” é composto por dados do tipo numérico com 12 caracteres. Os quatro primeiros caracteres se referem ao ano da notificação. O quinto e sexto dígitos são referentes ao mês da notificação e os outros dígitos são utilizados para referenciar a ordem de chegada das notificações, estando dispostos em ordem crescente. Esse campo é um campo de identificação única da notificação. A sua utilização na modelagem terá como objetivo manter a rastreabilidade das notificações. Ao analisar os registros presentes na base de dados, para esse campo, não foi identificado nenhum elemento faltante ou mesmo nenhum elemento que estivesse fora dos padrões esperados. Dessa maneira, não houve a necessidade de nenhuma preparação especial para os dados que compõem esse campo.

O campo “*tipo de notificação*” é do tipo *string* e tem como objetivo captar qual a natureza da notificação. Conforme visto anteriormente, existem apenas dois tipos de notificação em tecnovigilância: as Queixas Técnicas (QT) e os Eventos Adversos (EA). Por consequência, o presente campo possui apenas essas duas possibilidades. Ao analisar os

registros desse campo, verificou-se que não há nenhum registro faltante ou mesmo nenhum registro fora dos padrões esperados, não havendo, assim, a necessidade de tratamento específico para esses dados.

A “*descrição da ocorrência*” é do tipo *string*, sendo considerado o campo principal da modelagem. Esse campo representa a descrição objetiva sobre o EA ou QT, pelo notificante. A descrição é captada em texto livre, dessa maneira, não há qualquer tipo de controle, seja de tamanho, seja de formatação. Foi necessário utilizar alguns tratamentos prévios nos textos com a finalidade de normalizá-los e prepará-los para serem utilizados na fase de modelagem. Ao analisar os registros desse campo foi verificado que havia apenas um registro faltante, que será excluído da base de dados, uma vez que se trata de campo chave na modelagem dos dados para a extração das informações desejadas.

Por último, o quarto campo que será extraído da base de dados para ser utilizado na modelagem será o campo “*nome técnico do produto*”. Nele estão contidos todos os nomes técnicos dos produtos objetos das notificações. É um campo do tipo *string* e com entrada controlada, onde o notificador deve selecionar em uma lista o nome técnico correspondente ao produto objeto da notificação. O nome técnico do produto se refere a classe do produto. Alguns exemplos de nomes técnicos são “MARCA-PASSO INTRACARDÍACO” e “STENT PARA ARTÉRIAS CORONÁRIAS”. Ao analisar os registros contidos nesse campo, foi observado que existiam 5.602 dados faltantes. Esses dados serão substituídos pela palavra “*em branco*” para utilização na modelagem dos dados.

#### 4.4.2 Definição das classes e tags das Entidades Nomeadas

Antes de avançar para a parte de definição do modelo e dos atributos (*features*), que serão utilizados para auxiliar no Reconhecimento das Entidades Nomeadas, é necessário que se definam quais serão as entidades nomeadas que se deseja identificar. Para auxiliar nessa definição, serão apresentados a seguir alguns exemplos de textos, extraídos da base de dados. Foram retirados todos os dados que permitiriam a identificação do produto, do notificante ou do paciente:

##### **Exemplo 1**

“*Vazamento na conexão da extensão com o terminal*”

##### **Exemplo 2**

“*Dificuldade de manuseio devido qualidade do fio muito flexível*”

##### **Exemplo 3**

“*Introdutor com agulha romba não permitindo perfuração da pele outro produto também teve problema*”

Como o objetivo é identificar alterações nas frequências de EA ou QT, relacionadas aos dispositivos médicos, uma entidade que precisa ser identificada é a própria ocorrência do



EA ou QT. Assim, precisamos identificar as palavras ou conjunto de palavras (entidades), em cada um dos textos, que caracterizam esses **eventos**. No Exemplo 1, a palavra que caracteriza o **evento** ocorrido, no caso uma QT, é a ocorrência de um “vazamento”. Ao se analisar o Exemplo 2, é possível verificar que há um problema com relação a flexibilidade do fio. Nesse caso, o **evento** de interesse na notificação é a palavra “flexível”. No Exemplo 3, o **evento** de interesse é o fato da agulha ser “romba”. Assim, a entidade de interesse, segundo esses três exemplos, são as palavras relacionadas aos **eventos**. No caso dos três exemplos apresentados, todos eram relativos à ocorrência de um Queixa Técnica com o produto notificado.

A partir apenas dos **eventos** já seria possível identificar alterações nas frequências das QT ou EA relativos à um problema específico. Porém, é possível complementar essa informação com a identificação do **objeto** que guarda relação com o **evento** identificado. Após a identificação das palavras que caracterizam os eventos (as QT ou EA), no texto, é possível notar que as mesmas estão relacionados a algum objeto. No caso do Exemplo 1, o **evento** “vazamento” ocorreu com o **objeto** “conexão da extensão”. No Exemplo 2, o **evento** “flexível” está relacionado com o **objeto** “fio”. Já no Exemplo 3, o **evento** “romba” está relacionado como o **objeto** “agulha”.

A identificação do **evento** e do **objeto** relacionado a ele, é possível identificar e qualificar as alterações nas frequências de QT e EA classificadas como não graves. Dessa forma, no presente trabalho, as duas entidades de interesse serão a entidade **Evento** e a entidade **Objeto**. Caso seja possível identificar essas entidades nos textos, bastaria o cruzamento com os dados do registro do produto, que está contido nos formulários e, por consequência, na base de dados, para descobrir um aumento na frequência de determinado **Evento**, em um dado **Objeto**.

#### 4.4.3 Formação do Corpus

Com base na revisão do estado da arte, optou-se pela utilização de modelos baseados em aprendizado de máquina do tipo supervisionado, já que são os modelos mais utilizados e os que vêm apresentando os melhores resultados. Esses modelos exigem uma base de dados anotada para que possam ser treinados e testados. Essa base é conhecida como Corpus. No presente projeto, será realizada uma prova de conceito com a utilização de um tipo de notificação, referente a um tipo de produto.

A missão da Anvisa é “*promover a proteção da saúde da população, por intermédio do controle sanitário da produção e consumo de produtos e serviços submetidos à vigilância sanitária*”. Tendo em vista que a proteção da saúde da população é o objetivo da Anvisa, optou-se pela seleção, dentre os produtos mais notificados, do produto que expusesse seu usuário ao maior risco. Dentre todos os Dispositivos Médicos, os Dispositivos Médicos

Implantáveis (DMI) são os que implicam em maior risco para a saúde do seu usuário, por serem implantados no corpo. Dentre todos os DMI, o produto escolhido foi o Catéter, que é o DMI com o maior número de notificações, conforme mostra a Tabela 4.6. São 8.601 notificações no total, sendo que 7.477 notificações, ou 87% do total, são referentes à Queixas Técnicas. Assim, a prova de conceito será realizada com base nas notificações sobre Queixas Técnicas relativas aos Cateteres.

Uma questão importante, que deve ser levada em consideração, nesse caso, é que as notificações relativas a Cateter engloba uma séria de diferentes tipos de Cateteres, de diferentes naturezas e finalidades. Por exemplo, é possível encontrar desde cateter nasal até cateter venoso central. A separação entre esses tipos de cateter seria importante, uma vez que um evento como, por exemplo, o "rompimento", poderia implicar em maiores consequências se ocorresse com um cateter venoso central do que com um cateter jelco, por exemplo. Porém, Infelizmente, da forma como as notificações foram estruturadas, não é possível identificar o tipo de cateter que o produto faz parte. Essa é uma importante limitação do presente estudo.

O tamanho do GSC varia bastante na literatura. O corpus utilizado por Wunnava et al.[39] era composto por 1.089 Prontuários Eletrônicos. Três outros estudos utilizaram a base de dados n2c2, que é composta por 505 anotações clínicas extraídas da base de dados MIMIC-III [42, 41, 43]. Miftahutdinov et al.[36] utilizaram o corpus CADEC, composto por 1.250 relatos de eventos adversos. Chen et al.[34] selecionaram 1.129 mensagens das mídias sociais francesas que possuem pelo menos uma citação de eventos adversos. Molane-Hondere et al.[47] utilizaram 1.100 mensagens das mídias sociais francesas.

Assim, uma amostra de 1.700 notificações foi selecionada para a formação do corpus do presente trabalho, que é um número superior aos apresentados pelos estudos levantados. Essas notificações foram selecionadas por uma amostragem aleatória estratificada. Para que fosse possível conseguir uma boa representatividade na amostra, foi realizada uma clusterização de todas as notificações de QT sobre Cateteres. Para tal, as notificações foram transformadas em vetores numéricos através da utilização de *word embedding* pré-treinadas, com o objetivo de utilizar os clusters na estratificação da amostra. Ao todo, foram 125 clusters formados para as 7.477 notificações de QT sobre Cateteres. A clusterização foi realizada utilizando clusterização hierárquica.

Após a seleção das 1.706 notificações para a formação dos corpus, todas as notificações foram traduzidas para o inglês por meio do tradutor do Microsoft Word, que compõe o pacote Office 365. A tradução para o inglês tem como objetivo a realização de um experimento para comparação das performances dos diferentes modelos aplicados na base em português e na mesma base traduzida para o inglês. Os objetivos desse experimento são verificar o impacto da tradução e da utilização de *transfer-learning* no texto traduzido

no resultado final. Como a maior parte do mercado de dispositivos médicos brasileiro é formada por produtos importados, caso não haja uma diferença significativa entre os resultados da base traduzida por máquina e a base original em português, abre-se a possibilidade da utilização das bases de tecnovigilância de outros países para suplementar o sistema de tecnovigilância brasileira.

#### 4.4.4 Processo de Anotação

A anotação da base de dados foi realizada manualmente pelo autor, utilizando o sistema Microsoft EXCEL, do pacote Office 365. As notificações em português e as traduzidas para o inglês foram colocadas lado a lado, em colunas consecutivas. A anotação das 3.412 notificações, sendo 1.706 originais em português e as 1.706 notificações traduzidas por máquina para o inglês, foi realizada de maneira supervisionada.

Para que houvesse coerência nas anotações, elas foram realizadas por notificação. Assim, todas as palavras, ou *tokens*, das notificações em português eram classificadas e, em seguida, as palavras da mesma notificação, traduzida para o inglês, eram classificadas. A ideia foi garantir a coerência entre as classificações das mesmas entidades entre as diferentes línguas. Não houve nenhum tipo de correção ortográfica ou de qualquer outra natureza. Quando verificou-se que haviam erros ortográficos, ou mesmo erros de tradução, buscou-se uma interpretação sobre qual era a intenção do notificador. Ao final, foi realizada uma revisão de todas as entidades anotadas nos textos com o objetivo de se identificar eventuais inconsistências na base anotada. As entidades foram separadas por suas respectivas classificações e as inconsistências foram corrigidas.

Entre cada uma das palavras, ou *tokens*, foi inserido o marcador "<O>". Para que seja possível realizar os experimentos referentes à aplicação de diferentes representações de segmentos, optou-se por uma anotação inicial genérica que fosse facilmente transformada para as diferentes representações de segmento. A Tabela 4.9 apresenta os marcadores utilizados nas anotações iniciais.

Tabela 4.9: Tags e significados das Entidades Nomeadas

<b>Tags</b>	<b>Significado</b>
<2>	Início da Entidade <i>Evento</i>
<21>	Parte da Entidade <i>Evento</i>
<1>	Início da Entidade <i>Objeto</i>
<11>	Parte da Entidade <i>Objeto</i>
<O>	Não é Entidade

A anotação genérica facilita a substituição para diferentes representações de segmento. Por exemplo, a mesma tabela poderia ser reescrita para seguir o sistema de taguemento de

entidades nomeadas, desenvolvido por Ramshall et al.(1999)[72], a Tabela 4.10 apresenta como ficariam as cinco classes utilizando a representação de segmento BIO.

Tabela 4.10: Tags e significados das Entidades Nomeadas

<b>Tags</b>	<b>Significado</b>
B–eve	Início da Entidade <b><i>Evento</i></b>
I–eve	Parte da Entidade <b><i>Evento</i></b>
B–obj	Início da Entidade <b><i>Objeto</i></b>
I–obj	Parte da Entidade <b><i>Objeto</i></b>
O	Não é Entidade

Após todo o processo de anotação e verificação de inconsistências, a base de dados foi transformada para o formato de entrada dos modelos que foram desenvolvidos. Os dados foram organizados de maneira onde cada palavra fica alocada nas linhas e as informações sobre a identificação da notificação a qual a palavra pertence e a Tag da palavra ficam nas colunas. No caso das bases de dados utilizadas para o treinamento do CRF, foram acrescentadas mais duas outras colunas, uma referente a *POS-tagging* da palavra e outra coluna referente ao número de registro do produto notificado, já que ambos foram utilizados como *features* do modelo. Com o objetivo de não apenas validar a utilidade do resultado, mas também validar a consistência e utilidade do processo de anotação, o resultado final, incluindo estatísticas das entidades ***Evento*** e ***Objeto***, foram apresentados e validados para os técnicos da área responsável pela análise e investigação das notificações.

A Tabela 4.11 apresenta as estatísticas sobre o resultado obtido a partir da anotação do corpus. Considerando o corpus original, em português, foram anotadas 679 entidades ***Evento***, com uma frequência total de 4.431. No caso das entidades ***Objeto***, foram anotadas 194 entidades, com uma frequência total de 1.756.

Tabela 4.11: Estatísticas sobre as anotações realizadas no corpus

	<b>Treino</b>
Notificações	1.706
Entidades únicas de Eventos	679
Frequências entidades Eventos	4.431
Entidades únicas de Objetos	194
Frequências entidades Objetos	1.756

## 4.5 Modelagem

O objetivo do presente trabalho é desenvolver uma solução baseada em aprendizado de máquina que permita identificar alterações nas frequências de EA ou QT presentes no

sistema de notificação NOTIVISA. Dessa forma, será realizada uma prova de conceito baseada nas QT de Cateteres, por ser o DMI com o maior número de notificações. A partir do levantamento do estado da arte, é possível perceber que o NER é uma técnica de NLP que vem sendo amplamente aplicada para a detecção de diferentes Entidades de interesse em textos livres da área de saúde, principalmente, na detecção de EA.

O modelo *Conditional Random Fields* foi selecionado como linha de base do presente estudo. O modelo CRF foi escolhido por ser um modelo clássico de aprendizado de máquina que era considerado como o estado da arte antes da aplicação das redes neurais para o NER. O CRF será confrontado com um modelo de Redes Neurais Recorrentes de arquitetura BiLSTM + CRF, que pode ser considerado o estado da arte da aplicação de NER para vigilância pós-mercado de produtos de saúde. Ambos modelos utilizam diferentes tipos de *features*, que serão apresentadas a seguir.

#### 4.5.1 Atributos ou *features*

A maioria dos modelos de REN sustenta seus resultados, principalmente, em atributos (*features*) específicos escolhidos para rodar seus modelos. Com o auxílio desses atributos, seus modelos são capazes de captar as características das diferentes classes de suas entidades. A engenharia de atributos, que é o processo de se encontrar os atributos que melhor ajudam a diferenciar duas classes, ou entidades, de palavras, exige experimentos custosos de tentativa e erro[57].

Os atributos podem ser de diferentes naturezas. No caso do modelo de RNN aplicados para NER, os atributos mais utilizados são vetores de palavras, de caracteres ou mesmo que representam uma sentença. No caso do modelo CRF, Os atributos mais comuns são atributos relacionados ao próprio *token* (palavras) ou aos *tokens* (palavras) vizinhos. Alguns exemplos de atributos são:

- Prefixo ou sufixo da própria palavra;
- Prefixo ou sufixo das palavras vizinhas, podendo variar de 1 a n palavras anteriores e/ou posteriores;
- Atributos relacionados às expressões regulares;
- Atributos morfossintáticos como *POS-tagging*;
- *Stemming* da palavra e das palavras ao seu redor;

Além dos atributos, diversos outros fatores influenciam nos resultados obtidos pelos modelos de NER como a língua, o número de entidades extraídas, as técnicas utilizadas, os recursos disponíveis, os modelos de tags escolhidos e os atributos utilizados[14].

Choi et al.(2016)[73] combinaram em seu modelo os seguintes atributos: (1) formato da própria palavra, de duas palavras para trás e duas para frente; (2) *lemmatization* da própria palavra; (3) cluster da palavra e de duas palavras para frente; (4) o *word embendding* da própria palavra, de todas as quatro palavras para trás e de todas as quatro palavras para frente; (5) sufixo e prefixo da própria palavra; (6) *POS-tagging* da própria palavra, de duas para trás e duas para frente; (7) Tag das entidades nomeadas de três palavras para trás; (8) NER *gazeteer* da própria palavra, de todas as três palavras para trás e de todas as três palavras para frente e (9) o conjunto de recursos ortográficos.

Para a construção do modelo base, os atributos que foram testados, bem como o tipo de variável estão listados na Tabela 4.12:

Tabela 4.12: Atributos modelo base

<b>Atributo</b>	<b>Tipo</b>
Palavra em caixa baixa	<i>String</i>
As 2 últimas letras da palavra	<i>String</i>
As 3 últimas letras da palavra	<i>String</i>
<i>POS-tagging</i>	<i>String</i>
As 2 primeiras letras da palavra	<i>String</i>
As 3 primeiras letras da palavra	<i>String</i>
Número de registro do produto objeto da notificação	<i>Inteiro</i>
Se a primeira letra da palavra está em caixa alta	<i>Boolean</i>
Se a palavra é um número	<i>Boolean</i>
Stemming	<i>String</i>

Cada uma das *features* apresentadas na Tabela 4.12 foi testada em uma janela de até 3 palavras (tokens) para frente e para trás da entidade de interesse. As *POS-tagging* das palavras foram obtidas por meio de um modelo pré-treinado em bases de dados clínicas, em português, disponível na biblioteca FLAIR[74]. No caso dos *Stemming*, foi utilizada a função "*RSLP Portuguese stemmer*", da biblioteca NLTK.

No caso do modelo BiLSTM + CRF, a única *feature* utilizada foi a *word-embedding*, ou representações vetoriais de palavras. Essa foi a *feature* mais utilizada, com bons resultados, pelos artigos levantados na revisão de estado da arte. Serão testadas dois tipos de *word-embeddings*. Uma iniciada com valores randômicos e a outra proveniente de um modelo pré-treinado, com o objetivo de se verificar o impacto do *transfer-learning*.

#### 4.5.2 *Conditional Random Fields - CRF*

O CRF é um modelo utilizado para predições em cenários onde as informações de contexto influenciam no resultado da previsão. As séries temporais são um exemplo. Outro exemplo são os textos, que são formados por uma sequência de palavras que, em conjunto

e seguindo uma determinada ordem, dão um significado ao texto. Para a aplicação do CRF no presente trabalho, utilizou-se o estimador *crfsuite* da biblioteca *scikit-learn*. Esse estimador é baseado no *linear-chain CRF*.

O CRF modela a probabilidade condicional de  $P(Y/X)$ . No caso da aplicação do CRF em NER, a variável dependente  $Y$  se refere a Tag da entidade de interesse e a variável  $X$  é composta por um vetor de *features*  $X = (x_1, x_2, x_3...)$  composta por atributos da própria palavra e das palavras que as cercam.

Por levar em consideração o contexto, o CRF é um modelo capaz de identificar regras implícitas nas classificações de entidades. O modelo é capaz de aprender as a relação entre as tags e atributos das palavras de interesse, bem como das relações entre as tags e atributos das palavras próximas com a tag da palavra de interesse. Assim, o modelo é capaz de aprender as regras implícitas de classificação das entidades. Por exemplo, no caso da anotação BIO, o modelo é capaz de aprender que uma Tag B (*Begin* só pode vir no início de uma entidade e que a Tag I (*Inside* somente pode vir precedida de uma Tag B ou de outra Tag I).

Para a seleção das features e da janela de palavras, o corpus foi separado em uma base de treino e teste na razão de 80% e 20%. A partir dessa divisão, foram realizadas análises de sensibilidade para cada a inclusão dos parâmetros e para o aumento da janela de palavras, com o objetivo de identificar o efeito da inclusão ou exclusão de cada feature e da inclusão de mais palavras a esquerda e a direita da palavra de interesse no macro F1-score.

Dessa forma, o melhor resultado encontrado foi de uma janela de três palavras para a direita e três palavras para a esquerda e as features selecionadas foram as seguintes:

1. Número do registro do produto objeto da notificação;
2. *POS-tagging* da palavra e das três palavras da janela;
3. Palavra e as três palavras da janela em caixa baixa;
4. As três primeiras palavras da palavra;
5. As duas primeiras palavras da palavra;
6. Se a primeira letra da palavra é um número;
7. Se a primeira letrada da palavra e das palavras da janela estão em caixa alta.

### 4.5.3 Arquitetura BiLSTM + CRF

Conforme apresentado na Revisão do Estado da Arte, os modelos de Redes Neurais Recorrentes (RNN), com arquitetura BiLSTM + CRF são as mais utilizadas pela Literatura

nos últimos anos, para a aplicação de NER, sendo responsáveis pelos melhores resultados. No presente trabalho, essa arquitetura será confrontada com o modelo CRF. Os experimentos realizados serão descritos na seção 4.6.

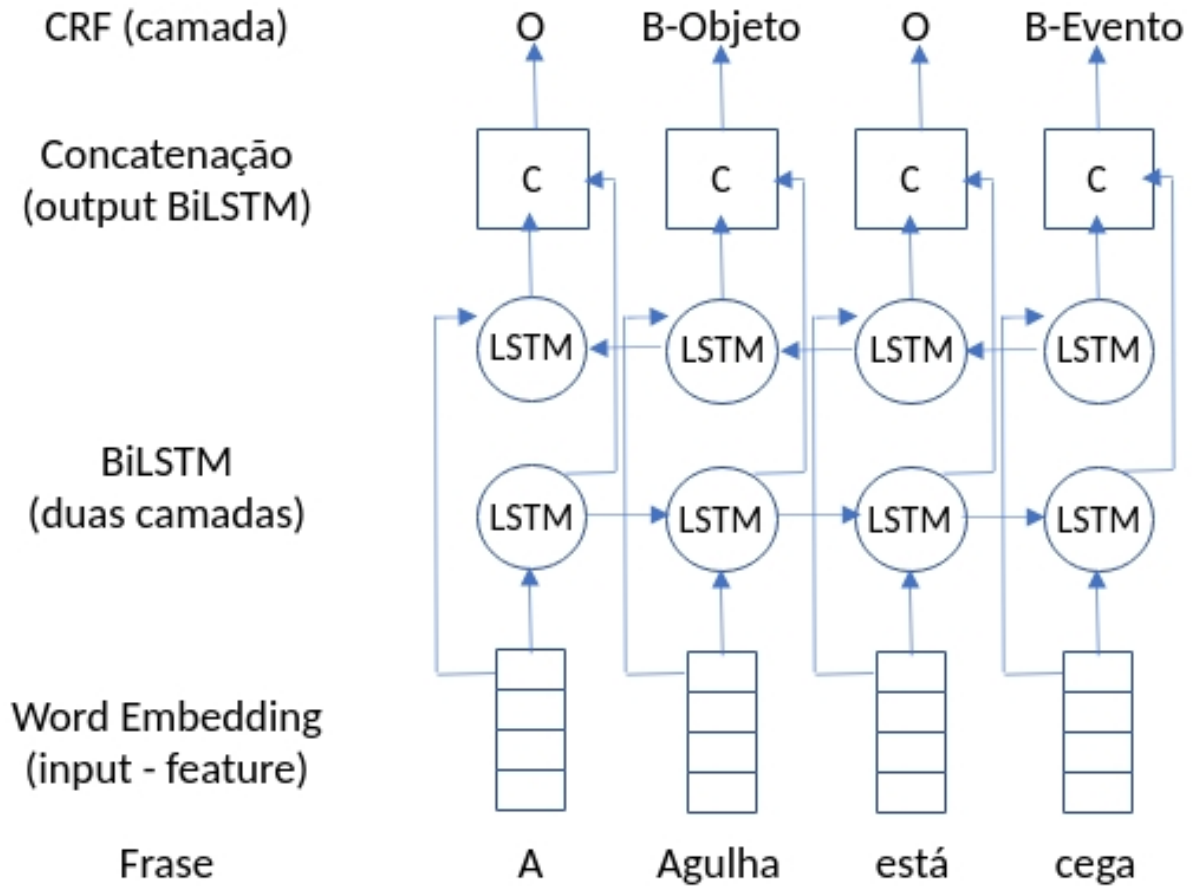


Figura 4.2: Modelo BiLSTM + CRF.

A Figura 4.2 apresenta uma representação da arquitetura BiLSTM + CRF utilizada no presente trabalho. O input do modelo são as *word embeddings*, representações numéricas das palavras que compõem a frase, no nosso caso, as notificações. Além de ser a *feature* mais utilizada na literatura, a opção pela *word embeddings* levou em consideração a variação no tamanho das notificações. O tamanho das notificações variam de uma única palavra até notificações com 40.000 caracteres. Esse cenário, por exemplo, não favorece a utilização de *sentence embeddings*. Dessa maneira, optou-se pela utilização das *word embeddings*.

As *word embeddings* são processadas por cada uma das camadas de LSTM. Uma das camadas processa considerando o contexto localizado a direita da palavra, enquanto a outra camada processa considerando o contexto localizado a esquerda da palavra de interesse. Os dois resultados dessas camadas são então concatenados em uma representação única.



Essa é uma representação dos contextos a direita e a esquerda da palavra de interesse e serve como input para a última camada CRF, que gera como resultado as classificações da palavra de interesse em uma das labels.

Além de *word embeddings* inicializadas com valores aleatórios, também foram utilizadas *word embeddings* pré-treinadas com o objetivo de verificar o impacto da transferência de conhecimento (*transfer learning*) no contexto de tecnovigilância. No caso das notificações em português, foi utilizada uma *word embedding* pré-treinada em uma base de dados composta por dezessete corpus diferentes, de fontes e gêneros diferentes, da língua portuguesa, contendo textos com português do Brasil e textos com português de Portugal[75]. No caso das notificações traduzidas para o inglês, as *word embeddings* foram do projeto GloVe, que utilizaram uma base de dados com 6 bilhões de tokens [76].

Para que o algoritmo utilizado na geração das *word embeddings* não influenciasse no resultado final, no caso da base em português, selecionou-se as *word embeddings* pré-treinadas com o algoritmo GloVe, assim como no caso da base em inglês. Da mesma maneira, para que o tamanho da *word embedding* não influenciasse nos experimentos, foram selecionadas *word embeddings* de tamanho 100 tanto para a base em português quanto para a base em inglês.

# Capítulo 5

## Resultados e Análises

Tendo em vista os objetivos geral e específicos, apresentados na Seção 1.3, foram realizados nove experimentos a consecução, totalizando 180 execuções dos algoritmos. Nesse capítulo serão apresentados os resultados alcançados em cada um desses experimentos. Para avaliar os resultados, foi realizada uma reamostragem por bootstrap para o corpus original em português e replicada para o corpus traduzido para o inglês. Dessa forma, as composições das 20 amostras de teste e treino ficaram idênticas tanto para a base original em português quanto para a base traduzida para o inglês.

As comparações entre os diferentes modelos foram realizadas por um teste T bicaudal pareado, utilizando uma distribuição t com 19 graus de liberdade e um nível de significância de 0,05. Os testes serão realizados com o objetivo de identificar se as diferenças entre os resultados do macro F1-score são estatisticamente significantes. Nesse caso, um valor-t acima de 2,0930 fará com que a hipótese nula de igualdade seja rejeitada, o que significa que a diferença encontrada para as macro f1-scores são estatisticamente significativas ao nível de significância de 0,05.

### 5.1 Escolha do modelo e *features*

Para que fosse possível desenvolver uma solução que permita a identificação das alterações nas frequências de eventos relacionados à cateteres, o primeiro passo foi a seleção do modelo com melhor desempenho para a identificação de entidades nomeadas. Nesse sentido, quatro experimentos foram realizados nas 20 amostras de treino e teste com o corpus original, em português. Os dois primeiros experimentos foram realizados com o objetivo de selecionar a melhor *feature* para o modelo BiLSTM + CRF, bem como para definir com quantas épocas o modelo seria treinado. Já os outros dois testes confrontaram o modelo BiLSTM + CRF com o modelo BiLSTM puro e com o modelo CRF.

### 5.1.1 Escolha das *features* e o impacto do *transfer learning*

Conforme definido na seção anterior, a única *feature* a ser utilizada no treinamento do BiLSTM + CRF será a word embedding, que foi a *feature* mais utilizada pelos estudos levantados na revisão de estado da arte. O vetor das word embeddings pode ser iniciado com valores randômicos ou pode ser inicializado com vetores numéricos pré-treinados em grandes bases de dados. Nesse último caso há uma "transferência do conhecimento", *transfer learning*.

Assim, o primeiro experimento terá como objetivo definir se há diferença entre os vetores pré-treinados e os vetores randômicos e, por consequência, definir se há impacto com o *transfer learning*. Para a condução do presente experimento, foram escolhidas as word embeddings pré-treinadas por Hartmann et al.[75], utilizando o algoritmo GloVe e com dimensão 100. A Tabela 5.1 apresenta os resultados do teste T pareado, indicando que há uma diferença estatisticamente significativa em favor do *transfer learning*. Ambos os modelos foram treinados com 50 épocas. Os resultados estão em linha com os estudos levantados que utilizavam word embeddings pré-treinadas como *feature*. A utilização de *transfer learning* gera resultados significativamente superiores quando comparado aos resultados obtidos com a inicialização com vetores randômicos.

Tabela 5.1: Comparação BiLSTM + CRF sem e com word embeddings pré-treinadas

	Valores
F1-score médio BiLSTM + CRF sem transfer learning	0,8455
F1-score médio BiLSTM + CRF com transfer learning	0,8730
Diferença	0,0275
IC de 95%	(0,0167;0,0383)
Valor-t	5,3397
p-valor	0,0000

### 5.1.2 BiLSTM + CRF 50 épocas vs BiLSTM + CRF 100 épocas

Considerando que o transfer learning demonstrou impactar a macro F1-score do modelo, o segundo experimento comparou o modelo BiLSTM + CRF treinado com 50 épocas e com 100 épocas, ambos utilizando word embeddings pré-treinadas. O objetivo era definir se haveria diferença entre o número de épocas no resultado final, já que quanto maior a quantidade de épocas, maior é o tempo exigido para o treinamento do modelo e maior é a quantidade de recursos. O resultado demonstrou uma diferença estatisticamente significativa em favor do treinamento com 100 épocas, conforme pode ser visto na Tabela 5.2. Assim, serão utilizadas 100 épocas, preferencialmente, para o treinamento do modelo BiLSTM + CRF.

Tabela 5.2: Comparação BiLSTM+CRF com treinamento de 50 e 100 épocas

	Valores
F1-score médio BiLSTM + CRF 100 épocas	0,8775
F1-score médio BiLSTM + CRF 50 épocas	0,8730
Diferença	0,0045
IC de 95%	(0,0005;0,0085)
Valor-t	2,3759
p-valor	0,0248

### 5.1.3 BiLSTM + CRF vs BiLSTM

O terceiro experimento comparou o modelo BiLSTM + CRF com o modelo BiLSTM puro, ambos com a utilização de word embeddings pré-treinadas e treinados com 100 épocas. Os resultados foram favoráveis ao modelo BiLSTM + CRF, conforme pode ser visto na Tabela 5.3. O objetivo aqui foi verificar o benefício da incorporação de uma última camada de CRF na arquitetura BiLSTM. A inclusão de uma última camada de CRF gerou um resultado significativamente superior ao apresentado pela arquitetura BiLSTM pura, o que vai ao encontro dos achados no levantamento de estado da arte, onde a arquitetura BiLSTM + CRF foi a predominante, principalmente, entre os estudos mais recentes.

Tabela 5.3: Comparação BiLSTM+CRF com BiLSTM

	Valores
F1-score médio BiLSTM + CRF	0,8775
F1-score médio BiLSTM	0,8365
Diferença	0,0410
IC de 95%	(0,0326;0,0494)
Valor-t	10,2465
p-valor	0,0000

### 5.1.4 BiLSTM + CRF vs CRF

O quarto e último experimento para a seleção do modelo a ser utilizado comparou o modelo BiLSTM + CRF com o modelo CRF. O modelo BiLSTM + CRF, para esse experimento, utilizou word embeddings pré-treinadas e foi treinado com 100 épocas. O resultado pode ser verificado na Tabela 5.4. O objetivo desse experimento foi confrontar o modelo considerado o estado da arte, BiLSTM + CRF, com o modelo base, nesse caso, o modelo CRF. O resultado não apontou para uma diferença estatisticamente significativa

para rejeitar a hipótese nula de igualdade entre os modelos. A causa para esse resultado é que em três amostras (amostra 3, amostra 4 e amostra 17) o modelo CRF não convergiu, gerando um resultado distorcido com f-scores de 0,18 para as três amostras, uma vez que houve uma sobreclassificação da classe predominante, no caso a classe "O". Assim, a média alcançada pelo CRF foi inferior à média esperada caso o modelo convergisse para as três amostras, fazendo com que o desvio-padrão fosse maior, o que gerou um intervalo de confiança mais amplo, contribuindo para a não rejeição da hipótese nula.

Dessa maneira, optou-se por rodar novamente o test t pareado excluindo-se os resultados das três amostras que não convergiram. Assim, o novo teste contou com 17 amostras e grau de liberdade 16. Nessas condições, a hipótese nula de igualdade será rejeitada para qualquer resultado do valor-t acima de 2,1200. O resultado desse novo teste está representado na Tabela 5.4. Retirando-se as três amostras em que o modelo CRF não convergiu, rejeita-se a hipótese nula, com o modelo BiLSTM + CRF apresentando um melhor resultado.

Tabela 5.4: Comparação BiLSTM+CRF com CRF para 20 amostras

	<b>Valores</b>
F1-score médio BiLSTM + CRF	0,8775
F1-score médio CRF	0,7665
Diferença	0,1110
IC de 95%	(-0,0090;0,2310)
Valor-t	1,9353
p-valor	0,0617

Tabela 5.5: Comparação BiLSTM+CRF com CRF para 17 amostras

	<b>Valores</b>
F1-score médio BiLSTM + CRF	0,8788
F1-score médio CRF	0,8700
Diferença	0,0088
IC de 95%	(0,0003;0,0174)
Valor-t	2,1845
p-valor	0,0387

## 5.2 Base em Português vs base traduzida para o Inglês

A partir dos primeiros experimentos, o modelo selecionado foi o BiLSTM + CRF com a utilização de word embeddings pré-treinadas. Esse modelo será utilizado como base para

realizar o experimento para comparar os desempenhos do modelo treinado e testado na base de dados original em português contra o mesmo modelo treinado e testado na mesma base de dados traduzida, por máquina, para o inglês. Esse teste é importante, pois se não houver uma diferença estatisticamente significativa entre as bases em português e inglês, haveria a possibilidade de utilizar bases em outras línguas, de outros países, para suplementar os sistemas de vigilância pós-mercado nacional.

As 20 amostras de treino e teste, das bases em português e traduzida para o inglês possuem as mesmas composições. Assim, é possível realizar o teste t pareado. Foram realizados dois experimentos e em todos esses experimentos, o modelo BiLSTM + CRF foi executado com 50 épocas. O primeiro experimento comparou as bases de dados utilizando word embedding com inicialização randômicas e o segundo experimento comparou as bases com a utilização de word embeddings pré-treinadas para ambas as línguas.

Os resultados dos dois experimentos estão descritos na Tabela Tabela 5.6, que diz respeito às bases sem pré-treinamento, e na Tabela 5.7, que traz os resultados do experimento realizado utilizando word embeddings pré-treinadas. Os resultados apontaram na mesma direção, ou seja, ao nível de significância de 0,05, não é possível rejeitar a hipótese nula de igualdade entre os modelos aplicados à base original em português e os mesmos modelos aplicados à base traduzida, por máquina, para o inglês. Esse é um resultado importante, uma vez que as novas tecnologias de saúde costumam ser importadas e, caso seja possível ter acesso a bases de notificações internacionais, essas poderiam ser utilizadas como suplementação do sistema nacional de vigilância pós-mercado, apontando Queixas Técnicas ou Eventos Adversos ocorridos em outros países que poderiam estar ocorrendo, ou viriam a ocorrer, no Brasil.

Tabela 5.6: Base português vs base traduzida para o inglês sem pré-treinamento

	<b>Valores</b>
F1-score médio base português	0,8455
F1-score médio base traduzida para o inglês	0,8375
Diferença	0,0080
IC de 95%	(-0,0017;0,0177)
Valor-t	1,7300
p-valor	0,0919

Tabela 5.7: Base português vs base traduzida para o inglês com pré-treinamento

	<b>Valores</b>
F1-score médio base português	0,8730
F1-score médio base traduzida para o inglês	0,8695
Diferença	0,0035
IC de 95%	(-0,0115;0,0045)
Valor-t	0,9174
p-valor	0,3584

### 5.3 Representação de segmento

A Seção 2.1.4 apresentou o resultado de dois estudos que realizaram experimentos sobre o impacto das diferentes representações de segmentos (RS) no resultado final dos modelos. O último experimento realizado teve como objetivo estudar esse impacto no resultado do presente estudo para definir a RS que será utilizada. Para tal, foram confrontadas três tipos de RS: BIO, BILOU e IOE.

A BIO foi selecionada por ser a mais utilizada na literatura levantada no presente estudo. A IOE foi selecionada por ser a RS com melhor desempenho segundo os experimentos realizados por Konkol et al.[52], que confrontaram a IOE com a BIO, BILOU e BIOE. A BILOU foi selecionada, pois foi apontada por Ratinov et al.[53] como a RS de melhor desempenho. Os experimentos foram conduzidos pela aplicação do teste T pareado, nas 20 amostras geradas a partir da base original em português, utilizando o modelo BiLSTM + CRF com word embedding pré-treinada em português e treinada com 100 épocas.

Os resultados apontaram que a BILOU foi superada tanto pela BIO, quanto pela IOE e que não é possível rejeitar a hipótese nula de igualdade entre a BIO e a IOE. Estes resultados são apresentados pela Tabela 5.8, pela Tabela 5.9 e pela Tabela 5.10, e vão de encontro aos achados de Ratinov et al.[53], onde a BILOU superou significativamente a BIO e estão parcialmente de acordo com os resultados de Konkol et al.[52], onde a IOE superou a BILOU e a BIO.

Tabela 5.8: BIO vs BILOU

	<b>Valores</b>
F1-score médio BIO	0,8775
F1-score médio BILOU	0,8430
Diferença	0,0345
IC de 95%	(0,0284;0,0406)
Valor-t	11,7826
p-valor	0,0000

Tabela 5.9: IOE vs BILOU

	<b>Valores</b>
F1-score médio IOE	0,8780
F1-score médio BILOU	0,8430
Diferença	0,0350
IC de 95%	(0,0291;0,0409)
Valor-t	12,3487
p-valor	0,0000

Tabela 5.10: BIO vs IOE

	<b>Valores</b>
F1-score médio BIO	0,8775
F1-score médio IOE	0,8780
Diferença	0,0005
IC de 95%	(-0,0024;0,0034)
Valor-t	0,3604
p-valor	0,7157

A partir dos achados nos experimentos realizados no presente estudo, o modelo selecionado foi o BiLSTM + CRF, treinado com 100 épocas e que utiliza como *feature* as word embeddings pré-treinadas. Outras conclusões é que as representações de segmento influenciam no resultado final. Porém, a representação de segmento ideal pode variar e deve ser definida caso a caso. Como os resultados obtidos pelas RS BIO e IOE são equivalentes, optou-se pela utilização da BIO, já que é a mais utilizada na literatura.



## 5.4 Aplicação do Modelo

A partir dos resultados dos experimentos apresentados acima, o modelo selecionado para aplicação em toda a base de dados de notificações de cateter foi o modelo do tipo supervisionado, com arquitetura BiLSTM + CRF, treinado com 100 épocas, aplicado ao corpus em português anotado utilizando a representação de segmento BIO. Para avaliar a aplicação prática do modelo, o mesmo foi treinado em todo o corpus, 1.706 notificações de cateter, e testado em todas as outras notificações referentes a cateter. Como existem muitos erros de classificação do tipo de notificação, QT ou EA, por parte dos notificantes, optou-se pela aplicação do modelo em toda a base de notificações de cateteres, não apenas para as notificações classificadas como QT.

Tabela 5.11: Estatísticas sobre a base de treino e teste

	<b>Treino</b>	<b>Teste</b>	<b>Total</b>
Notificações	1.706	6.895	8.601
Entidades únicas de Eventos	679	1.995	2.243
Frequências entidades Eventos	4.431	9.296	13.727
Entidades únicas de Objetos	194	818	873
Frequências entidades Objetos	1.756	6.737	8.493

A Tabela 5.11 apresenta o consolidado do resultado obtido com a aplicação do modelo a toda a base de dados. O corpus anotado de referência, composto por 1.706 notificações, foi utilizado como base de treino do modelo, que foi testado no restante da base de dados de notificações de cateter, ou seja, nas outras 6.895 notificações, totalizando as 8.601 notificações.

Analisando a base de treino, existiam 679 e 194 entidades únicas referentes às entidades Eventos e Objetos, respectivamente. O modelo foi capaz de encontrar 1.995 e 818 entidades únicas referentes às entidades Eventos e Objetos, respectivamente, na base de teste. Além das classificações equivocadas, a maior parte das entidades classificadas pelo modelo na base de treino era referente à erros ortográficos, abreviação de palavras, variações da mesma palavra (ex: conjugação verbal) e sinônimos. Essas capacidades apresentadas pelos modelo são importantes para sinalizar a capacidade do modelo de generalizar.

Tabela 5.12: Estatísticas sobre as Entidades classificadas pelo modelo

	<b>Quantidade</b>	<b>%</b>
Entidades únicas <i>Evento</i> presente na base de <i>treino</i> e não classificadas na base de <i>teste</i>	87	10%
Entidades únicas <i>Evento</i> classificadas na base de <i>teste</i> e não presentes na base de <i>treino</i>	1.564	78%
Entidades únicas <i>Objeto</i> presente na base de <i>treino</i> e não classificadas na base de <i>teste</i>	22	11%
Entidades únicas <i>Objeto</i> classificadas na base de <i>teste</i> e não presentes na base de <i>treino</i>	679	83%

As estatísticas sobre as entidades classificadas estão listadas na Tabela 5.12. Dentre as Entidades *Evento* e *Objeto* classificadas, o modelo não foi capaz de classificar 87 e 22 entidades referentes à *Evento* e *Objeto*, respectivamente, que não foram classificadas na base de treino, mas estavam presentes na base de teste. Essas entidades que estavam presentes na base de treino e teste, mas que o modelo não foi capaz de classificá-las eram todas com baixa frequência. Tendo em vista que o objetivo do presente estudo é identificar alterações nas frequências das notificações, a não identificação de entidades pouco frequentes não compromete a utilização do modelo.

Por outro lado, o modelo apresentou uma boa capacidade de identificação de entidades que não estavam presentes na base de treino. Ao todo, foram encontradas 1.564 e 679 entidades referentes à *Evento* e *Objeto*, respectivamente, que foram classificadas na base de teste e não estavam presentes na base de treino. Dentre esses, existem alguns erros de classificação, porém, em sua grande maioria, dizem respeito à erros ortográficos, variações na escrita, abreviação de palavras e sinônimos.

A Tabela 5.13 apresenta uma primeira análise manual sobre os classificações realizadas pelo modelo na base de teste. Foram realizadas análises sob três perspectivas: (1) classificações corretas, na base de teste, das entidades que estavam presentes na base de treino, (2) classificações corretas, na base de teste, das entidades que não estavam presentes na base de treino e (3) classificações equivocadas, na base de teste, das entidades que não estavam presentes na base de treino.

Pela análise preliminar, o modelo conseguiu bons resultados tanto na classificação das entidades *Evento*, quanto das entidades *Objeto*. De todas as entidades classificadas, 82,16% das entidades *Evento* foram corretamente classificadas, sendo que 11,10% das entidades corretamente classificadas foram referentes a novas entidades encontradas pelo modelo. Por outro lado, o modelo equivocou-se na classificação de 17,84% das entidades.

Já no caso da classificação das entidades *Objeto*, o modelo alcançou 87,24%, onde apenas 3,61% eram correspondentes a novas entidades, que não estavam presentes na base de treino. Por outro lado, 12,77% das entidades foram classificadas erroneamente pelo modelo e 83,63% das entidades classificadas estavam presentes na base de testes. O modelo apresentou um bom desempenho seja na classificação geral das entidades, seja na identificação de novas entidades.

Tabela 5.13: Estatísticas sobre as classificações das entidades

	<b>Objeto</b>	<b>%</b>	<b>Evento</b>	<b>%</b>
Frequência entidades corretamente classificada na base de teste que estava presente na base de treino	5.634	83,63%	6.592	71,06%
Frequência entidades corretamente classificada na base de teste que não estava presente na base de treino	243	3,61%	1.030	11,10%
Frequência entidades equivocadamente classificada na base de teste que não estava presente na base de treino	860	12,77%	1.655	17,84%
<b>TOTAL</b>	<b>6.737</b>	<b>100,00%</b>	<b>9.277</b>	<b>100,00%</b>

O resultado das classificações foi compilado em um painel de BI, elaborado utilizando o software Microsoft Power BI. O referido painel foi apresentado e validado por especialistas da área de tecnovigilância com o objetivo de verificar a utilidade dos resultados e o impacto potencial do trabalho para a suplementação do sistema tradicional de vigilância pós-mercado. A Figura 5.1 apresenta o referido painel. Os dados referentes ao número de registro foram ocultados, já que os dados de notificação são confidenciais. A figura mostra que foram classificadas 8.447 notificações, de um total de 8.601 notificações existentes. Ao todo, 154 notificações não tiveram nenhuma entidades de interesse, Evento ou Objeto, classificada pelo modelo.

O painel possui ao todo doze campos, sendo seis campos de pesquisa e outros seis campos de resultado. Por meio do painel, é possível buscar por número de registro, ano da notificação, mês da notificação, tipo de notificação (QT ou EA), além das pesquisas por evento e objeto, que são compostos pelas entidades classificadas pelo modelo. Com o objetivo de facilitar a busca e tornar a ferramenta mais precisa, após a classificação das entidades, foram criados dicionários para os dois tipos de entidades. Os objetivos foram agrupar erros ortográficos, variações da mesma palavras (ex: conjugações verbais) e agrupar sinônimos. No caso da entidade de interesse *Evento*, as 2243 entidades deu

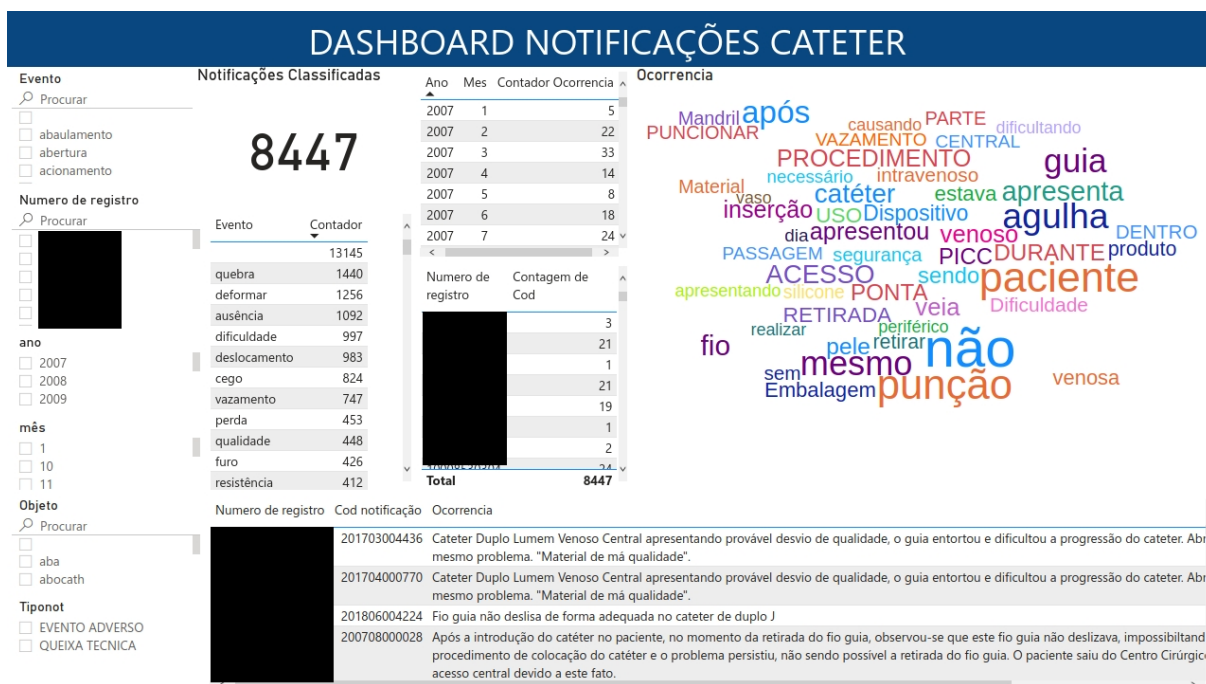


Figura 5.1: Painel de BI com os resultados da aplicação do modelo em toda a base de notificações de cateter.

origem a um dicionário com 172 entidades, já para a entidade de interesse *Objeto*, as 818 entidades deu origem a um dicionário com 68 entidades.

Os outros seis campos referem-se aos resultados. Um apresenta o numero de notificações com pelo menos uma entidade de interesse classificadas pelo modelo, Objeto ou Evento. Outro campo de resultado diz respeito a uma nuvem de palavras para auxiliar o usuário na extração de informação. Existem quatro campos do tipo tabela. Um deles lista os eventos e a quantidade de notificações que apresentaram o registro daquele evento. Outro apresenta a frequência de notificação por mês e ano, outro apresenta a quantidade de notificações por registro e a tabela de baixo apresenta o número de registro, o código da notificação e o texto contendo o relato da notificação.

Antes mesmo de se aplicar filtros, o painel já oferece algumas informações importantes. Além do número de notificações, é possível saber qual o mês e o ano que tiveram a maior frequência de notificações, no caso, foram os meses de setembro e outubro de 2018. É possível também saber quais os produtos são responsáveis pelo maior número de notificações. No caso, o produto responsável pelo maior número de notificações foi responsável, sozinho, por 713 notificações. Outra curiosidade é que, dentre todos os 482 produtos que tiveram pelo menos uma notificação no período, os 10 primeiros são responsáveis por, aproximadamente, 39% do total de notificações.

Outro dado que é possível extrair do painel, sem a realização de qualquer filtro é a

relação das entidades de interesse *Evento* mais frequentes. Dentre as cinco primeiras, é possível destacar a ocorrência de quebras, deformação, a ausência de algum item, alguma dificuldade em relação ao uso do produto e problemas com o deslocamento. Importante ressaltar que as 10 entidades de interesse *Evento* mais frequentes são responsáveis por, aproximadamente, 66% do total de notificações.

#### 5.4.1 Detecção de alterações nas frequências de Queixas Técnicas de Cateter

A missão da ANVISA é "*Proteger e promover a saúde da população, mediante a intervenção nos riscos decorrentes da produção e do uso de produtos e serviços sujeitos à vigilância sanitária, em ação coordenada e integrada no âmbito do Sistema Único de Saúde*". Existem riscos decorrentes do uso de Dispositivos Médicos (DM) e o papel da tecnovigilância, ou vigilância pós-mercado de DM, é identificar possíveis Queixas Técnicas (QT) ou Eventos Adversos (EA) relacionados ao uso desses DM para que seja possível uma ação tempestiva com a finalidade de se evitar maiores danos à saúde da população que seria exposta ao DM.

Dessa forma, o objetivo principal do presente estudo é a captação de alterações nas frequências de ocorrências de QT. Três dimensões são importantes para a captação dessa alteração de frequência. A primeira é o produto, que pode ser pesquisado pelo seu número de registro. A segunda é o *Evento* e o terceiro é o tempo. Por exemplo, seria possível descobrir a frequência de notificações de um tipo de *Evento* para um determinado produto, filtrando pelo tipo de *Evento* e pelo número de registro do produto. Para descobrir alterações nas frequências, bastaria verificar os dados referentes as ocorrências nos meses e anos.

O resultado da classificação permite ao órgão regulador realizar a pesquisa sob a perspectiva que melhor lhe atende. Por exemplo, o órgão pode decidir por priorizar as análises pelos *Eventos* que mais incorrem em risco. Por exemplo, o *Evento* "Quebra" pode incorrer em maior risco que o *Evento* "Ausência" de alguma parte do DM, já que o rompimento de um cateter que encontra-se dentro de um paciente pode provocar danos mais sérios do que a ausência de uma parte de um cateter, que inviabiliza o seu uso. Assim, o órgão pode priorizar o acompanhamento dos *Eventos* "Quebra" em detrimento do acompanhamento dos dos *Eventos* "Ausência"

Para exemplificar, foi realizado um levantamento sobre o comportamento e conteúdo das notificações relacionadas ao dos *Eventos* "Quebra". Dessa forma, foi utilizado o filtro de *Eventos* para selecionar o dos *Evento* "Quebra". A partir desse filtro, é possível obter os dados sobre os produtos que mais tiveram notificações dessa natureza. No caso, o

primeiro colocado teve 107 notificações referentes ao dos *Evento* "Quebra". Ao analisar o comportamento da frequência, foi possível perceber que a média é de uma notificação e meia por mês. Nesse período, foi possível identificar dois picos, um em outubro de 2015 e outro em dezembro de 2016, ambos com 9 notificações cada, sendo que uma alteração já havia sido percebida em setembro de 2015 com 4 notificações desse dos *Eventos*. Ao se filtrar pelo número do registro e pelos meses e ano de interesse, chegou-se ao conjunto de notificações.

Após a leitura, é possível notar que os eventos estão relacionados ao rompimento do cateter. Assim, a Agência poderia agir junto ao fabricante para compreender esses eventos e entender se os casos são isolados, se estariam relacionados a um lote específico e o que poderia ser adotado para que se evitasse ou minimizasse o risco de ocorrências desse tipo de *Evento*. Caso houvesse a possibilidade de distinguir entre os diferentes tipos de cateteres, a ação seria ainda mais eficiente, uma vez que seria possível pesquisar pelo tipo de cateter e perceber diferenças de ocorrência do *Evento* "Quebra" com os cateteres menos invasivos, como por exemplo os cateteres nasais e os cateteres jelco, bem como com os mais invasivos, como os cateteres venosos centrais. A quebra desse último teria o potencial maior de ocasionar danos mais graves à saúde do que a quebra de um cateter jelco. Como dito anteriormente, essa é uma limitação do presente estudo.

Outro exemplo, agora com o *Evento* "Deslocamento", foi possível notar picos consideráveis com 4 produtos diferentes. No primeiro produto, as notificações são bem distribuídas ao longo do tempo, porém, em janeiro de 2014 houve um pico com 11 notificações, todas sinalizando para problemas relacionados a maleabilidade do cateter, o que atrapalhava o seu deslocamento. Dois outros casos chamaram a atenção. Em um deles, o produto teve "apenas" 52 notificações. Destas, 45 se concentraram nos meses de agosto e setembro de 2010 e se referiam a problemas com a dobra dos cateteres, transfixação, ponta cega e lubrificação, o que atrapalhava o deslocamento do cateter. Outro caso na mesma linha, aconteceu com outro produto, que teve 28 das suas 34 notificações no período, concentrados no mês de novembro de 2015. Ao se analisar as notificações, percebeu-se que todas elas se referiam a um problema relacionado ao bisel cego, que não permitia o deslocamento do cateter.

Ao se analisar os dados por eventos e ao se analisar os 10 produtos com maiores números de notificações, foi possível notar dois tipos de comportamentos das distribuições de frequências das QT. Um grupo de produtos possui frequências relativamente baixas, porém, constantes ao longo do tempo. Outro grupo de produtos possui frequências esporádicas com picos maiores em curtos espaços de tempo. Esses tipos de comportamento poderão exigir diferentes tipos de ação por parte do órgão regulador. Por exemplo, os casos em que os eventos são constantes ao longo do tempo podem sinalizar um problema

crônico com o produto, o que demandaria uma ação diferente dos casos esporádicos, mas de alta frequência, que poderiam sinalizar um problema relacionado à um determinado lote do produto, o que demandaria uma ação rápida do órgão regulador para a retirada ou suspensão da comercialização do lote, se fosse o caso.

A partir da classificação das entidades de interesse, pela aplicação do modelo de NER, não apenas é possível verificar alterações nas frequências, mas é possível que se faça um mapeamento "epidemiológico" do mercado de cateteres, descobrindo quais os produtos estão relacionados com determinado tipo de *Eventos*, quais são os tipos de *Eventos* que devem ser priorizados de acordo com os riscos que estão envolvidos, quais são os eventos mais frequentes, quais os riscos inerentes a cada um dos produtos monitorados, quais os tipos de *Eventos* mais frequentes para cada tipo de produto entre outras análises possíveis. Essas ações, que até então não eram conhecidas, já que estavam armazenadas em textos livres, podem, inclusive, alterar as ações dos órgãos reguladores, que agora podem tomar suas decisões com base em dados e evidências.

# Capítulo 6

## Considerações Finais

No presente trabalho foi desenvolvida uma solução baseada em aprendizado de máquina para extração de informações, da base de dados de notificações pós-mercado de dispositivos médicos, com o objetivo de viabilizar a identificação de alterações nas frequências de eventos relacionados às notificações classificadas como não graves. Essas alterações nas frequências permitem a Agência agir de forma tempestiva para evitar ou minimizar maiores danos à saúde da população que seria exposta ao produto.

Para tal, foi realizada uma prova de conceito com as notificações relativas à Catecter. Foi desenvolvido um modelo supervisionado para o Reconhecimento de Entidades Nomeadas (NER) com a utilização do modelo BiLSTM + CRF, treinado com 100 épocas e tendo como *feature* word embeddings pr-treinadas [75]. A opção por esse modelo, com essa configuração, se deu a partir da realização de experimentos que confrontaram o modelo BiLSTM + CRF com os modelos BiLSTM e CRF isoladamente, confrontaram o modelo BiLSTM + CRF treinado com 50 e 100 épocas e confrontaram o modelo BiLSTM + CRF utilizando-se de *word embedding* pré-treinadas e *word embeddings* randômicas. Esse modelo alcançou um F1-score de 87,75%

O corpus foi anotado utilizando a representação de segmento BIO, que é a mais utilizada na literatura e a que apresentou um resultado estatisticamente significativo quando comparada ao BILOU e apresentou resultado equivalente quando comparada à IOE. Esse é um achado importante, pois corrobora com alguns achados na literatura sobre o impacto da utilização de diferentes representações de segmento no resultado final dos modelos. Confrontando os achados do presente estudo com outros achados na literatura, não é possível saber qual a melhor representação de segmento. Assim, quando for aplicar o NER, é importante testar diferentes representações de segmento para identificar qual oferece o melhor resultado no caso específico.

Outros importante achado do presente estudo é a possibilidade de utilização de bases de dados em outras línguas para suplementar o sistema nacional de vigilância pós-



mercado. A partir da realização de dois experimentos, que confrontaram os resultados obtidos na aplicação do modelo selecionado à base de dados em português, com e sem pré-treinamento, e na aplicação do modelo selecionado à base de dados traduzida por máquina para o inglês, com e sem pré-treinamento, não foi possível rejeitar a hipótese nula de igualdade entre os modelos. Esse é um importante achado, uma vez que o mercado de dispositivos médicos é dominado por produtos importados, assim, esse resultado abre caminho para a utilização de bases de dados em outras línguas para complementar as informações do sistema nacional de vigilância pós-mercado. Estudos futuros seriam necessários para verificar a possibilidade de utilizar a base traduzida por máquina para o inglês como base de treinamento e uma base original, em inglês, como base de teste.

O modelo apresentou um bom resultado no reconhecimento de entidades nomeadas quando aplicado a toda a base de dados de cateter. Apresentou uma boa capacidade na identificação de novas entidades como sinônimos, erros ortográficos, palavras abreviadas e variações da mesma palavra. Esse resultado gera boas expectativas para a utilização do modelo nas bases de dados de outros produtos. Essa é uma limitação do estudo, já que foi aplicado apenas a base de dados de cateter, impossibilitando a generalização dos resultados aqui alcançados para outras bases de dados, de outros modelos. O treinamento do modelo aqui apresentado para as bases de dados de outros produtos é uma oportunidade para estudos futuros. Outra importante limitação, já discutida ao longo do texto, diz respeito ao fato das notificações de Cateteres englobarem uma série de diferentes cateteres, não sendo possível a sua estratificação.

Por fim, foi elaborado um painel de Business Intelligence (BI) com vistas a permitir a extração de informações por parte dos usuários de maneira simples, rápida e intuitiva. A partir do painel, foi possível obter informações quanto aos produtos mais notificados, os tipos de eventos mais frequentes, incluindo uma análise por produto e, principalmente, permite a identificação de alterações nas frequências de eventos, seja em uma busca genérica, por todo o mercado, seja uma alteração de frequência nos eventos relacionados à determinados produtos.

# Referências

- [1] Organization, World Health *et al.*: *Global atlas of medical devices*. 2017. 1
- [2] (Anvisa), Agência Nacional de Vigilância Sanitária: *Resolução nº 185, de 13 de outubro de 2006*. Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF, Seção 1:25–25, nov 2001. 1, 2
- [3] Majjers, Maria C e Francisus B Niessen: *Prevalence of rupture in poly implant prothese silicone breast implants, recalled from the european market in 2010*. Plastic and reconstructive surgery, 129(6):1372–1378, 2012. 2
- [4] (Anvisa), Agência Nacional de Vigilância Sanitária: *Tecnovigilância - Alerta 1015*. <http://www.anvisa.gov.br/sistec/alerta/RelatorioAlerta.asp?NomeColuna=CO{ }SEQ{ }ALERTA{&}Parametro=1015>, acesso em 31 de março de 2019. 2
- [5] Rathi, Vinay K. e Stacey T. Gray: *Device Safety*. Otolaryngologic Clinics of North America, 52(1):103–114, feb 2019, ISSN 00306665. <https://doi.org/10.1016/j.otc.2018.08.013><https://linkinghub.elsevier.com/retrieve/pii/S0030666518301683>. 3
- [6] Zuckerman, Diana, Paul Brown e Aditi Das: *Lack of Publicly Available Scientific Evidence on the Safety and Effectiveness of Implanted Medical Devices*. JAMA Internal Medicine, 174(11):1781, nov 2014, ISSN 2168-6106. <http://archinte.jamanetwork.com/article.aspx?doi=10.1001/jamainternmed.2014.4193>. 3
- [7] Dhruva, Sanket S, Lisa A Bero e Rita F Redberg: *Strength of study evidence examined by the fda in premarket approval of cardiovascular devices*. Jama, 302(24):2679–2685, 2009. 3
- [8] Wirth, Rüdiger e Jochen Hipp: *Crisp-dm: Towards a standard process model for data mining*. Em *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*, páginas 29–39. Citeseer, 2000. 3, 8, 23, 24, 33
- [9] Djuriscic, Snezana, Ana Rath, Sabrina Gaber, Silvio Garattini, Vittorio Bertele, Sandra Nadia Ngwabyt, Virginie Hivert, Edmund AM Neugebauer, Martine Laville, Michael Hiesmayr *et al.*: *Barriers to the conduct of randomised clinical trials within all disease areas*. Trials, 18(1):360, 2017. 3, 34

- [10] Cheng, Shunfeng, Diganta Das e Michael G Pecht: *Using failure modes, mechanisms, and effects analysis in medical device adverse event investigations*. Em *ICBO*, 2011. 3
- [11] MARIANO, Ari Melo e Maíra Santos ROCHA: *Revisão da literatura: Apresentação de uma abordagem integradora*. Em *XXVI Congreso Internacional de la Academia Europea de Dirección y Economía de la Empresa (AEDEM)*, Reggio Calabria, volume 26, 2017. 9
- [12] Harzing, Anne Wil e Satu Alakangas: *Google scholar, scopus and the web of science: a longitudinal and cross-disciplinary comparison*. *Scientometrics*, 106(2):787–804, 2016. 10
- [13] Grishman, Ralph e Beth M Sundheim: *Message understanding conference-6: A brief history*. Em *COLING 1996 Volume 1: The 16th International Conference on Computational Linguistics*, 1996. 12
- [14] Goyal, Archana, Vishal Gupta e Manish Kumar: *Recent named entity recognition and classification techniques: a systematic review*. *Computer Science Review*, 29:21–43, 2018. 12, 15, 16, 18, 19, 20, 21, 28, 29, 48
- [15] Galper, Benjamin Z, David E Beery, Gregory Leighton e Libbe L Englander: *Comparison of adverse event and device problem rates for transcatheter aortic valve replacement and mitralclip procedures as reported by the transcatheter valve therapy registry and the food and drug administration postmarket surveillance data*. *American heart journal*, 198:64–74, 2018. 13, 17
- [16] Callahan, Alison, Jason A Fries, Christopher Ré, James I Huddleston, Nicholas J Giori, Scott Delp e Nigam H Shah: *Medical device surveillance with electronic health records*. *NPJ digital medicine*, 2(1):1–10, 2019. 13, 15, 18, 21
- [17] Toledo, J Ignacio, Manuel Carbonell, Alicia Fornés e Josep Lladós: *Information extraction from historical handwritten document images with a context-aware neural model*. *Pattern Recognition*, 86:27–36, 2019. 15, 17, 19
- [18] Liu, Xiaojun, Ning Yang, Yu Jiang, Lichuan Gu e Xianzhang Shi: *A parallel computing-based deep attention model for named entity recognition*. *The Journal of Supercomputing*, 76(2):814–830, 2020. 15, 17, 19, 20
- [19] Thomas, Anu e S Sangeetha: *An innovative hybrid approach for extracting named entities from unstructured text data*. *Computational Intelligence*, 35(4):799–826, 2019. 15, 17, 18, 20
- [20] Liao, Fei, Liangli Ma, Jingjing Pei e Linshan Tan: *Combined self-attention mechanism for chinese named entity recognition in military*. *Future Internet*, 11(8):180, 2019. 15, 17, 19, 20
- [21] AP, Ajees, Sumam Mary Idicula *et al.*: *An improved word representation for deep learning based ner in indian languages*. *Information*, 10(6):186, 2019. 15, 17, 19

- [22] Shardlow, Matthew, Meizhi Ju, Maolin Li, Christian O'Reilly, Elisabetta Iavarone, John McNaught e Sophia Ananiadou: *A text mining pipeline using active and deep learning aimed at curating information in computational neuroscience*. *Neuroinformatics*, 17(3):391–406, 2019. 15, 17, 18, 20
- [23] Qin, Ya, Guo wei Shen, Wen bo Zhao, Yan ping Chen, Miao Yu e Xin Jin: *A network security entity recognition method based on feature template and cnn-bilstm-crf*. *Frontiers of Information Technology & Electronic Engineering*, 20(6):872–884, 2019. 15, 17, 19, 20
- [24] Xu, Kai, Zhenguo Yang, Peipei Kang, Qi Wang e Wenyin Liu: *Document-level attention-based bilstm-crf incorporating disease dictionary for disease named entity recognition*. *Computers in biology and medicine*, 108:122–132, 2019. 15, 17, 18, 19, 20
- [25] Xie, Zhe, Yuanyuan Yang, Mingqing Wang, Ming Li, Haozhe Huang, Dezhong Zheng, Rong Shu e Tonghui Ling: *Introducing information extraction to radiology information systems to improve the efficiency on reading reports*. *Methods of information in medicine*, 58(02/03):094–106, 2019. 15, 17, 19, 20
- [26] Zhang, Xiaohui, Yaoyun Zhang, Qin Zhang, Yuankai Ren, Tinglin Qiu, Jianhui Ma e Qiang Sun: *Extracting comprehensive clinical information for breast cancer using deep learning methods*. *International Journal of Medical Informatics*, 132:103985, 2019. 15, 17, 19, 20
- [27] Weber, Leon, Jannes Münchmeyer, Tim Rocktäschel, Maryam Habibi e Ulf Leser: *Huner: improving biomedical ner with pretraining*. *Bioinformatics*, 36(1):295–302, 2020. 15, 16, 17, 21
- [28] Shi, Xue, Yingping Yi, Ying Xiong, Buzhou Tang, Qingcai Chen, Xiaolong Wang, Zongcheng Ji, Yaoyun Zhang e Hua Xu: *Extracting entities with attributes in clinical text via joint deep learning*. *Journal of the American Medical Informatics Association*, 26(12):1584–1591, 2019. 15, 17
- [29] El-allaly, Ed drissiya, Mourad Sarrouiti, Nouredine En-Nahnahi e Said Ouatik El Alaoui: *An adverse drug effect mentions extraction method based on weighted online recurrent extreme learning machine*. *Computer methods and programs in biomedicine*, 176:33–41, 2019. 15, 17, 20
- [30] Kim, Juae, Youngjoong Ko e Jungyun Seo: *A bootstrapping approach with crf and deep learning models for improving the biomedical named entity recognition in multi-domains*. *IEEE Access*, 7:70308–70318, 2019. 15, 17, 20
- [31] Wei, Hao, Mingyuan Gao, Ai Zhou, Fei Chen, Wen Qu, Chunli Wang e Mingyu Lu: *Named entity recognition from biomedical texts using a fusion attention-based bilstm-crf*. *IEEE Access*, 7:73627–73636, 2019. 15, 17, 20
- [32] Sun, Lin, Fule Ji, Kai Zhang e Chi Wang: *Multilayer toi detection approach for nested ner*. *IEEE Access*, 7:186600–186608, 2019. 15, 17, 21

- [33] Xia, Long, G Alan Wang e Weiguo Fan: *A deep learning based named entity recognition approach for adverse drug events identification and extraction in health social media*. Em *International Conference on Smart Health*, páginas 237–248. Springer, 2017. 15, 17, 19
- [34] Chena, Xiaoyi, Myrtille Deldossib, Rim Aboukhamisc, Carole Faviezd, Badisse Dahamnae, Pierre Karapetiantza, Armelle Guenegou-Arnouxa, Yannick Girardeauh, Sylvie Guillemin-Lanneb, Agnès Lillo-Le-Louët *et al.*: *Mining adverse drug reactions in social media with named entity recognition and semantic methods*. Em *MEDINFO 2017: Precision Healthcare Through Informatics: Proceedings of the 16th World Congress on Medical and Health Informatics*, volume 245, página 322. IOS Press, 2018. 15, 18, 19, 21, 45
- [35] Chen, Xiaoyi, Carole Faviez, Stéphane Schuck, Agnès Lillo-Le-Louët, Nathalie Texier, Badisse Dahamna, Charles Huot, Pierre Foulquié, Suzanne Pereira, Vincent Leroux *et al.*: *Mining patients' narratives in social media for pharmacovigilance: adverse effects and misuse of methylphenidate*. *Frontiers in pharmacology*, 9:541, 2018. 15, 18, 19
- [36] Miftahutdinov, Zulfat e Elena Tutubalina: *End-to-end deep framework for disease named entity recognition using social media data*. Em *2017 IEEE 30th Neumann Colloquium (NC)*, páginas 000047–000052. IEEE, 2017. 15, 17, 20, 45
- [37] Chapman, Alec B, Kelly S Peterson, Patrick R Alba, Scott L DuVall e Olga V Patterson: *Detecting adverse drug events with rapidly trained classification models*. *Drug safety*, 42(1):147–156, 2019. 15, 17, 19
- [38] Chen, Yao, Changjiang Zhou, Tianxin Li, Hong Wu, Xia Zhao, Kai Ye e Jun Liao: *Named entity recognition from chinese adverse drug event reports with lexical feature based bilstm-crf and tri-training*. *Journal of biomedical informatics*, 96:103252, 2019. 15, 16, 17, 19, 20, 21
- [39] Wunnava, Susmitha, Xiao Qin, Tabassum Kakar, Cansu Sen, Elke A Rundensteiner e Xiangnan Kong: *Adverse drug event detection from electronic health records using hierarchical recurrent neural networks with dual-level embedding*. *Drug safety*, 42(1):113–122, 2019. 15, 17, 22, 45
- [40] Zhang, Shinan, Shantanu Dev, Joseph Voyles e Anand S Rao: *Attention-based multi-task learning in pharmacovigilance*. Em *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, páginas 2324–22328. IEEE, 2018. 15, 17, 20
- [41] Dandala, Bharath, Venkata Joopudi, Ching Huei Tsou, Jennifer J Liang e Parthasarathy Suryanarayanan: *Extraction of information related to drug safety surveillance from electronic health record notes: Joint modeling of entities and relations using knowledge-aware neural attentive models*. *JMIR medical informatics*, 8(7):e18417, 2020. 15, 17, 20, 21, 45

- [42] Wei, Qiang, Zongcheng Ji, Zhiheng Li, Jingcheng Du, Jingqi Wang, Jun Xu, Yang Xiang, Firat Tiryaki, Stephen Wu, Yaoyun Zhang *et al.*: *A study of deep learning approaches for medication and adverse drug event extraction from clinical text*. Journal of the American Medical Informatics Association, 27(1):13–21, 2020. 15, 17, 20, 21, 45
- [43] Yang, Xi, Jiang Bian, Ruogu Fang, Ragnhildur I Bjarnadottir, William R Hogan e Yonghui Wu: *Identifying relations of medications with adverse drug events using recurrent convolutional neural networks and gradient boosting*. Journal of the American Medical Informatics Association, 27(1):65–72, 2020. 15, 17, 20, 21, 45
- [44] Christopoulou, Fenia, Thy Thy Tran, Sunil Kumar Sahu, Makoto Miwa e Sophia Ananiadou: *Adverse drug events and medication relation extraction in electronic health records with ensemble deep learning methods*. Journal of the American Medical Informatics Association, 27(1):39–46, 2020. 15, 17, 20, 21
- [45] Zhang, Tongxuan, Hongfei Lin, Yuqi Ren, Zhihao Yang, Jian Wang, Shaowu Zhang, Bo Xu e Xiaodong Duan: *Adversarial transfer network with bilinear attention for the detection of adverse drug reactions from social media*. Applied Soft Computing, 106:107358, 2021. 15, 17, 20
- [46] Zhang, Tongxuan, Hongfei Lin, Yuqi Ren, Zhihao Yang, Jian Wang, Xiaodong Duan e Bo Xu: *Identifying adverse drug reaction entities from social media with adversarial transfer learning model*. Neurocomputing, 453:254–262, 2021. 15, 17, 20
- [47] Morlane-Hondère, François, Cyril Grouin e Pierre Zweigenbaum: *Identification of drug-related medical conditions in social media*. Em *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC’16)*, páginas 2022–2028, 2016. 15, 16, 19, 21, 45
- [48] Wunnava, Susmitha, Xiao Qin, Tabassum Kakar, Xiangnan Kong, Elke A Rundensteiner, Sanjay K Sahoo e Suranjan De: *One size does not fit all: An ensemble approach towards information extraction from adverse drug event narratives*. Em *HEALTHINF*, páginas 176–188, 2018. 15, 16, 17, 19
- [49] Suárez-Paniagua, Víctor, Renzo M Rivera Zavala, Isabel Segura-Bedmar e Paloma Martínez: *A two-stage deep learning approach for extracting entities and relationships from medical texts*. Journal of biomedical informatics, 99:103285, 2019. 15, 17, 19, 20
- [50] Sun, Wencheng, Zhiping Cai, Yangyang Li, Fang Liu, Shengqun Fang e Guoyan Wang: *Data processing and text mining technologies on electronic medical records: a review*. Journal of healthcare engineering, 2018, 2018. 15
- [51] Kanimozhi, U e D Manjula: *A systematic review on biomedical named entity recognition*. Em *International Conference on Data Science Analytics and Applications*, páginas 19–37. Springer, 2017. 15, 16

- [52] Konkol, Michal e Miloslav Konopík: *Segment representations in named entity recognition*. Em *International Conference on Text, Speech, and Dialogue*, páginas 61–70. Springer, 2015. 20, 58
- [53] Ratinov, Lev e Dan Roth: *Design challenges and misconceptions in named entity recognition*. Em *Proceedings of the Thirteenth Conference on Computational Natural Language Learning (CoNLL-2009)*, páginas 147–155, 2009. 20, 58
- [54] Wissler, Lars, Mohammed Almashraee, Dagmar Monett Díaz e Adrian Paschke: *The gold standard in corpus annotation*. Em *IEEE GSC*, 2014. 21
- [55] Nadkarni, Prakash M, Lucila Ohno-Machado e Wendy W Chapman: *Natural language processing: an introduction*. Journal of the American Medical Informatics Association, 18(5):544–551, 2011. 25, 26, 31
- [56] Gupta, Vishal, Gurpreet S Lehal *et al.*: *A survey of text mining techniques and applications*. Journal of emerging technologies in web intelligence, 1(1):60–76, 2009. 25, 26
- [57] Habibi, Maryam, Leon Weber, Mariana Neves, David Luis Wiegandt e Ulf Leser: *Deep learning with word embeddings improves biomedical named entity recognition*. Bioinformatics, 33(14):i37–i48, 2017. 27, 31, 48
- [58] Màrquez, Lluís e Horacio Rodríguez: *Part-of-speech tagging using decision trees*. Em *European Conference on Machine Learning*, páginas 25–36. Springer, 1998. 28
- [59] Greene, Barbara B e Gerald M Rubin: *Automatic grammatical tagging of English*. Department of Linguistics, Brown University, 1971. 28
- [60] Pota, Marco, Fiammetta Marulli, Massimo Esposito, Giuseppe De Pietro e Hamido Fujita: *Multilingual pos tagging by a composite deep architecture based on character-level features and on-the-fly enriched word embeddings*. Knowledge-Based Systems, 164:309–323, 2019. 28
- [61] Kohavi, Ron *et al.*: *A study of cross-validation and bootstrap for accuracy estimation and model selection*. Em *Ijcai*, volume 14, páginas 1137–1145. Montreal, Canada, 1995. 29
- [62] Lafferty, John, Andrew McCallum e Fernando CN Pereira: *Conditional random fields: Probabilistic models for segmenting and labeling sequence data*. 2001. 31
- [63] Sutton, Charles, Andrew McCallum *et al.*: *An introduction to conditional random fields*. Foundations and Trends® in Machine Learning, 4(4):267–373, 2012. 31
- [64] Hochreiter, Sepp e Jürgen Schmidhuber: *Long short-term memory*. Neural computation, 9(8):1735–1780, 1997. 32
- [65] Dulleck, Uwe e Rudolf Kerschbamer: *On doctors, mechanics, and computer specialists: The economics of credence goods*. Journal of Economic literature, 44(1):5–42, 2006. 34

- [66] Marcus, Hani J., Christopher J Payne, Archie Hughes-Hallett, Adam P Marcus, Guang Zhong Yang, Ara Darzi e Dipankar Nandi: *Regulatory approval of new medical devices: Cross sectional study*. BMJ (Online), 353:i2587, may 2016, ISSN 17561833. <https://www.bmj.com/content/353/bmj.i2587.full>. 34
- [67] Garattini, Silvio, Janus C Jakobsen, Jørn Wetterslev, Vittorio Bertele, Rita Banzi, Ana Rath, Edmund AM Neugebauer, Martine Laville, Yvonne Masson, Virginie Hivert *et al.*: *Evidence-based clinical practice: overview of threats to the validity of evidence and how to minimise them*. European Journal of Internal Medicine, 32:13–21, 2016. 34
- [68] Bernard, Alain, Michel Vaneau, Isabelle Fournel, Hubert Galmiche, Patrice Nony e Jean Michel Dubernard: *Methodological choices for the clinical development of medical devices*. Medical Devices (Auckland, NZ), 7:325, 2014. 34
- [69] Casal, Roberto F., Lara Bashoura, David Ost, Hsienchang T. Chiu, Saadia A. Faiz, Carlos A. Jimenez, Rodolfo C. Morice e George A. Eapen: *Detecting Medical Device Complications*. American Journal of Medical Quality, 28(1):69–75, jan 2013, ISSN 1062-8606. <http://journals.sagepub.com/doi/10.1177/1062860612449475>. 35
- [70] (Anvisa), Agência Nacional de Vigilância Sanitária: *Resolução nº 67, de 21 de Dezembro de 2009*. Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF, Seção 1, dec 2009. 35, 36
- [71] Brasil: *Manual de tecnovigilância: abordagens de vigilância sanitária de produtos para a saúde comercializados no brasil*, 2010. 36, 40
- [72] Ramshaw, Lance A e Mitchell P Marcus: *Text chunking using transformation-based learning*. Em *Natural language processing using very large corpora*, páginas 157–176. Springer, 1999. 47
- [73] Choi, Jinho D: *Dynamic feature induction: The last gist to the state-of-the-art*. Em *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, páginas 271–281, 2016. 49
- [74] Akbik, Alan, Duncan Blythe e Roland Vollgraf: *Contextual string embeddings for sequence labeling*. Em *COLING 2018, 27th International Conference on Computational Linguistics*, páginas 1638–1649, 2018. 49
- [75] Hartmann, Nathan, Erick Fonseca, Christopher Shulby, Marcos Treviso, Jessica Rodrigues e Sandra Aluisio: *Portuguese word embeddings: Evaluating on word analogies and natural language tasks*. arXiv preprint arXiv:1708.06025, 2017. 52, 54, 67
- [76] Pennington, Jeffrey, Richard Socher e Christopher D Manning: *Glove: Global vectors for word representation*. Em *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, páginas 1532–1543, 2014. 52



## Anexo I

# Formulário para Notificação de Evento Adverso associado ao uso de Artigo Médico-Hospitalar

# Manual do Usuário



Notivisa – Sistema Nacional de Notificações  
para a Vigilância Sanitária

## Anexo 8

**Formulário para Notificação de Evento Adverso associado ao  
uso de Artigo Médico-Hospitalar**

## Sumário

1. Formulário para Notificação de Evento Adverso associado ao uso de Artigo Médico-Hospitalar.....	3
<b>1.1. ABA MOTIVO .....</b>	<b>3</b>
1.1.1. Bloco 4 – Evento Adverso .....	3
<b>1.2. ABA PRODUTO E EMPRESA .....</b>	<b>5</b>
1.2.1. Bloco 5 – Produto e empresa.....	5
1.2.2. Bloco 6 – Dados do Produto .....	6
1.2.3. Bloco 7– Dados do Fabricante (seleção da opção “Não” no campo 6.9).8	
1.2.4. Bloco 7– Dados do Importador (seleção da opção “Sim” no campo 6.9).9	
<b>1.3. ABA PACIENTE .....</b>	<b>10</b>
1.3.1. Bloco 8 – Informações do paciente ou usuário .....	10
1.3.2. Bloco 9 – Evolução do caso .....	13
<b>1.4. ABA OUTRAS INFORMAÇÕES.....</b>	<b>14</b>
<b>1.5. ABA PENDÊNCIAS .....</b>	<b>18</b>
1.5.1. Bloco – Pendências .....	18

## 1. Formulário para Notificação de Evento Adverso associado ao uso de Artigo Médico-Hospitalar

Para Evento Adverso de Artigo Médico Hospitalar (implante ortopédico, DIU, válvula cardíaca, seringa, catéter, equipo, etc.) estão disponíveis as seguintes abas:

- Motivo;
- Produto e Empresa;
- Paciente;
- Outras informações;
- Pendências.

### 1.1. Aba Motivo

Formulário para notificação de Evento Adverso associado ao uso de Artigo Médico-Hospitalar  
Número da Notificação: 2006.10.002190

Motivo	Produto e Empresa	Paciente	Outras Informações	Pendências
--------	-------------------	----------	--------------------	------------

#### 1.1.1. Bloco 4 – Evento Adverso

Este Bloco contém os campos para descrição e detalhamento do Evento Adverso.

4 - Evento Adverso		
4.1. Descreva detalhadamente o Evento Adverso : (*)		
<input type="text"/>		
4.2. Selecione o evento adverso (WHO-ART):		
<input type="text"/>	<input type="button" value="Pesquisar"/>	
4.3. Data da ocorrência: (*)		
<input type="text" value="Dia"/>	<input type="text" value="Mês"/>	<input type="text" value="Ano"/>
4.4. Local da ocorrência:		4.5. Endereço do local da ocorrência:
<input type="text" value="Estabelecimento de saúde"/>		<input type="text" value="515N - Edifício ANVISA"/>
4.4.1. Nome do estabelecimento de saúde:		
<input type="text" value="Empresa de Teste Ltda"/>		
4.4.2. CNPJ do estabelecimento de saúde:		
<input type="text" value="33.683.202/0001-34"/>		
4.6. País: (*)	4.7. UF: (*)	4.8. Município: (*)
<input type="text" value="BRASIL"/>	<input type="text" value="Distrito Federal - DF"/>	<input type="text" value="BRASÍLIA"/>

**Campo 4.1. Descreva detalhadamente o Evento Adverso:**

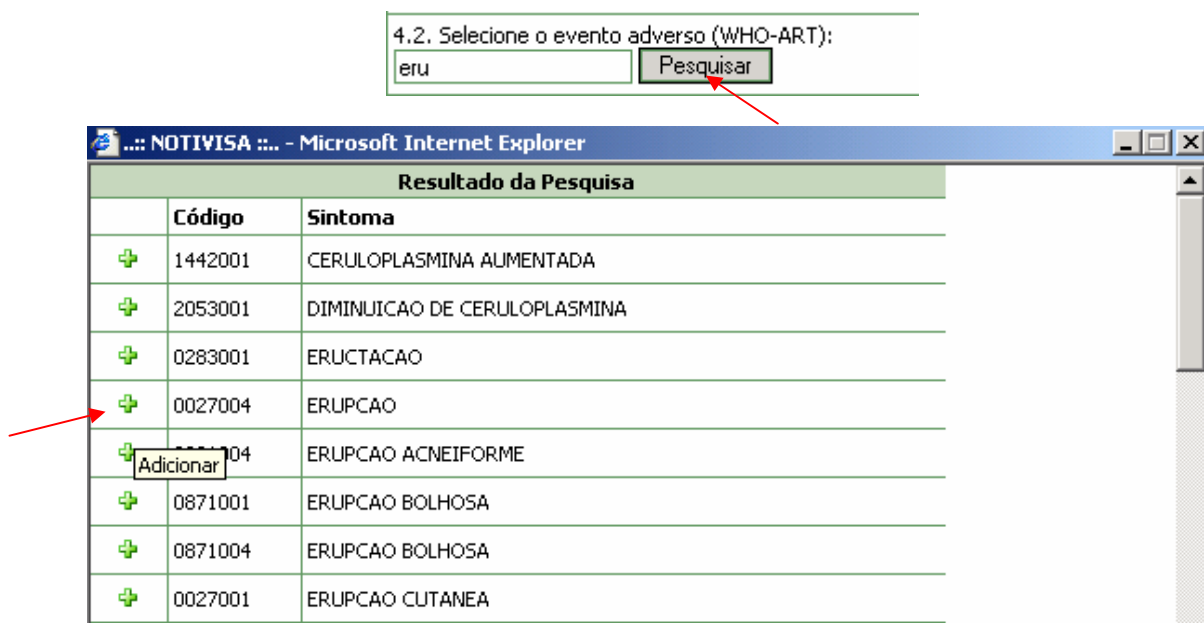
Campo de preenchimento obrigatório. Descreva detalhadamente o evento adverso incluindo os motivos da notificação.

**Campo 4.2. Selecione o evento adverso (WHO-ART):**

Campo para pesquisa da terminologia padronizada de evento adverso no WHO-ART (World Health Organization – Adverse Reaction Terminology).

Para a realização desta pesquisa digite, pelo menos, 3 caracteres do sinal/sintoma a ser pesquisado e clique em “pesquisar”.

Após aberta a tela *pop-up* com as opções, selecione a opção desejada clicando no símbolo **+**, localizado à esquerda da tela. Veja o exemplo:



Caso queira selecionar outro sinal/sintoma, repita a operação.

**Campo 4.3. Data da ocorrência:**

Campo de preenchimento obrigatório. Informe a data da ocorrência do evento adverso.

**Campo 4.4. Local da ocorrência:**

Selecione o local da ocorrência do evento adverso dentre as seguintes opções:

- Estabelecimento de Saúde;
- Residência;
- Outros.

Caso selecione a opção “Estabelecimento de Saúde”, preencha os seguintes campos:

4.4.1. Nome do estabelecimento de saúde:

4.4.2. CNPJ do estabelecimento de saúde:

Caso selecione a opção “Outro”, preencha o seguinte campo:

4.4.3. Citar:

#### Campo 4.5. Endereço do local da ocorrência:

Informe o endereço do local da ocorrência do evento adverso.

#### Campo 4.6. País:

Campo de preenchimento obrigatório. Informe o país onde ocorreu o evento adverso. A opção “Brasil” está pré-selecionada.

#### Campo 4.7. UF:

Informe o estado onde ocorreu o evento adverso. Este campo deve ser selecionado antes da seleção do campo 4.8. *Município*. Se o país selecionado no campo 4.6 for “Brasil”, este campo é de preenchimento obrigatório.

#### Campo 4.8. Município:

Informe o município onde ocorreu o evento adverso. É possível selecionar este campo somente após a seleção do campo 4.7. *UF*. Se o país selecionado no campo 4.6 for “Brasil”, este campo é de preenchimento obrigatório.

## 1.2. Aba Produto e Empresa

Nesta Aba, estão disponíveis os seguintes Blocos: **Produto e Empresa**, **Dados do Produto** e **Dados do Fabricante ou Importador**.



### 1.2.1. Bloco 5 – Produto e empresa

Este Bloco contém os campos para preenchimento do número do registro ou cadastro na ANVISA/MS e do CNPJ da empresa fabricante ou importadora. Para facilitar o preenchimento de dados desta aba, opcionalmente, o notificador poderá realizar a pesquisa no banco de dados da Anvisa.

5 - Produto e Empresa	
5.1. Número do registro ou cadastro na ANVISA/MS: <input type="text"/> <input type="button" value="Pesquisar"/>	5.2. CNPJ da empresa fabricante ou importador: <input type="text"/> <input type="button" value="Pesquisar"/>

Dependendo da pesquisa realizada, os campos do Bloco 5 – **Produto e Empresa**, Bloco 6 - **Dados do produto** e Bloco 7 – **Dados do Fabricante ou Importador** poderão ser preenchidos automaticamente.

Estes campos podem ser preenchidos automaticamente por pesquisa realizada em campos do Bloco 6 – **Dados do Produto**.

#### Campo 5.1. Número do registro na ANVISA/MS:

Campo numérico com espaço para até 13 dígitos. Informe o número do registro do produto na ANVISA/MS, impresso no rótulo/embalagem do produto.

Após preenchimento, clique no botão “pesquisar”.

Caso os dados do produto sejam encontrados no banco de dados da Anvisa, alguns campos referentes a esse produto e empresa serão preenchidos automaticamente.

Estes dados devem ser verificados e, caso não confiram com o produto associado à notificação do evento adverso, devem ser alterados manualmente pelo notificador.

#### Campo 5.2. CNPJ da empresa fabricante ou importador:

Campo numérico com espaço para até 14 dígitos. Informe o CNPJ da empresa fabricante ou importadora impresso no rótulo/embalagem do produto.

Após preenchimento, clique no botão “pesquisar”.

Caso os dados da empresa sejam encontrados no banco de dados da Anvisa, alguns campos referentes ao produto e empresa serão preenchidos automaticamente.

Estes dados devem ser verificados e, caso não confiram com o produto associado à notificação do evento adverso, devem ser alterados manualmente pelo notificador.

### 1.2.2. Bloco 6 – Dados do Produto


Este Bloco contém os campos para o preenchimento dos dados referentes ao produto, que poderão ser preenchidos automaticamente como resultado da pesquisa realizada no Bloco 5.

6 - Dados do Produto	
6.1. Nome comercial do produto: (*)	
<input type="text"/>	<input type="button" value="Pesquisar"/>
6.2. Produto:	
<input type="text" value="...:Selecione:..."/>	
6.3. Data de fabricação:	6.4. Data de validade:
<input type="text" value="Dia"/> <input type="text" value="Mês"/> <input type="text" value="Ano"/>	<input type="text" value="Dia"/> <input type="text" value="Mês"/> <input type="text" value="Ano"/>
6.5. Data de esterilização:	6.6. Data de validade da esterilização:
<input type="text" value="Dia"/> <input type="text" value="Mês"/> <input type="text" value="Ano"/>	<input type="text" value="Dia"/> <input type="text" value="Mês"/> <input type="text" value="Ano"/>
6.7. Número do lote/série:	6.8. O produto foi reprocessado?
<input type="text"/>	<input type="radio"/> Sim <input type="radio"/> Não
6.9. Produto importado?	
<input type="radio"/> Sim <input type="radio"/> Não	
<input type="button" value="Limpar dados da empresa e do produto"/>	

**Campo 6.1. Nome comercial do produto:**

Campo de preenchimento obrigatório para identificação do produto. Preencha com o nome comercial do produto constante em seu rótulo/embalagem.

Para a realização da pesquisa a partir deste campo, pelo menos 3 (três) caracteres devem ser informados.

Após clicar em “Pesquisar”, selecione uma das opções da lista apresentada. Caso os dados não confirmem com o produto em mãos, clique no botão  e preencha todos os campos manualmente.

Caso tenha sido realizada pesquisa pelo CNPJ, no campo 5.2, o campo 6.1 poderá apresentar lista com produtos fabricados pela empresa em questão para seleção do produto que motivou o evento adverso.

**Campo 6.2. Produto:**

Campo para seleção da categoria do produto. Este campo não pode ser preenchido manualmente.

Após a realização de pesquisa nos campos 5.1., 5.2. ou 6.1., o sistema poderá carregar automaticamente uma tabela para a seleção da categoria do produto. Caso não haja opções para seleção, desconsidere este campo.

**Campo 6.3. Data de fabricação:**

Selecione a data de fabricação do produto.

**Campo 6.4. Data de validade:**

Selecione a data de validade do produto.

**Campo 6.5. Data de esterilização:**

Selecione a data de esterilização do produto.

**Campo 6.6. Data de validade da esterilização:**

Selecione a data de validade da esterilização do produto.

**Campo 6.7. Número do lote/série:**

Campo alfa-numérico, de tamanho igual a 14. Informe o número do lote/série do produto.

**Campo 6.8. O produto foi reprocessado?**

Selecione a opção “Sim” caso o produto tenha sido reprocessado. Nesse caso preencha o seguinte campo:

**Campo 6.8.1. Quantas vezes foi reprocessado antes do incidente?**

Selecione a opção “Não” caso o produto não tenha sido reprocessado.

**Campo 6.9. Produto importado?**

Selecione a opção “Sim” caso o produto seja importado. Esta informação pode ser verificada no rótulo/embalagem do produto.



Selecione a opção “Não” caso o produto seja nacional. Esta informação pode ser verificada no rótulo/embalagem do produto.

### 1.2.3. Bloco 7– Dados do Fabricante (seleção da opção “Não” no campo 6.9)

Este Bloco contém os campos para preenchimento dos dados do fabricante.

Caso tenha sido realizada pesquisa nos campos anteriores, os campos deste bloco poderão ser preenchidos automaticamente.

6.8. Produto importado? <input type="radio"/> Sim <input checked="" type="radio"/> Não	
<b>7 - Dados do fabricante</b>	
7.1. Nome ou razão social do fabricante: (*) <input type="text"/>	
7.2. Endereço do fabricante: <input type="text"/>	
7.3. Número do telefone/SAC do fabricante: <input type="text"/>	
7.4. UF do fabricante: <input type="text" value="...: Selecione ..."/>	7.5. Município do fabricante: <input type="text" value="...: Selecione ..."/>
<input type="button" value="Limpar dados da empresa e do produto"/>	

#### Campo 7.1. Nome ou razão social do fabricante:

Campo de preenchimento obrigatório. Informe o nome ou razão social do fabricante que consta no rótulo/embalagem do produto.

#### Campo 7.2. Endereço do Fabricante:

Informe o endereço do fabricante que consta no rótulo/embalagem do produto.

#### Campo 7.3. Número do telefone/SAC do fabricante:

Informe o número do telefone ou Serviço de Atendimento ao Cliente (SAC) do fabricante que consta no rótulo/embalagem do produto.

#### Campo 7.4. U.F. do fabricante:

Selecione o estado do fabricante que consta no rótulo/embalagem do produto. Este campo deve ser selecionado antes da seleção do campo 7.5. *Município do fabricante.*

#### Campo 7.5. Município do fabricante:

Selecione o município do fabricante que consta no rótulo/embalagem do produto. Apenas é possível selecionar este campo após a seleção do campo 7.4. *UF do fabricante.*

### 1.2.4. Bloco 7– Dados do Importador (seleção da opção “Sim” no campo 6.9)

Este Bloco contém os campos para preenchimento dos dados do importador.

Caso tenha sido realizada pesquisa nos campos anteriores, os campos deste bloco poderão ser preenchidos automaticamente.

6.8. Produto importado? <input checked="" type="radio"/> Sim <input type="radio"/> Não	
<b>7 - Dados do importador</b>	
7.1. Nome ou razão social do importador: (*) <input type="text"/>	
7.2. Endereço do importador: <input type="text"/>	
7.3. Número do telefone/SAC do importador: <input type="text"/>	
7.4. UF do importador: ...: Selecione ... <input type="button" value="v"/>	7.5. Município do importador: ...: Selecione ... <input type="button" value="v"/>
7.6. Nome do fabricante: <input type="text"/>	7.7. País do fabricante: ...: Selecione ... <input type="button" value="v"/>
<input type="button" value="Limpar dados da empresa e do produto"/>	

#### Campo 7.1. Nome ou razão social do importador:

Campo de preenchimento obrigatório. Informe o nome ou razão social do importador que consta no rótulo/embalagem do produto.

#### Campo 7.2. Endereço do Importador:

Informe o endereço do importador que consta no rótulo/embalagem do produto.

#### Campo 7.3. Número do telefone/SAC do importador:

Informe o número do telefone ou Serviço de Atendimento ao Cliente (SAC) do importador que consta no rótulo/embalagem do produto.

#### Campo 7.4. U.F. do importador:

Selecione o estado do importador que consta no rótulo/embalagem do produto. Este campo deve ser selecionado antes da seleção do campo 7.5. *Município do importador.*

#### Campo 7.5. Município do importador:

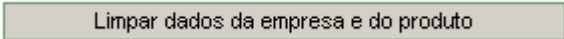
Selecione o município do importador que consta no rótulo/embalagem do produto. Apenas é possível selecionar este campo após a seleção do campo 7.4. *UF do importador.*

**Campo 7.6. Nome do fabricante:**

Informe o nome ou razão social do fabricante que consta no rótulo/embalagem do produto.

**Campo 7.7. País do Fabricante:**

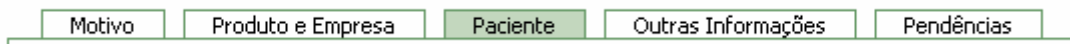
Selecione o país do fabricante do produto que consta no rótulo/embalagem.

**Botão Limpar dados da empresa e do produto**

Caso haja inconsistência ou incoerência no resultado da pesquisa realizada, clique nesse botão e os dados preenchidos nesta aba serão limpos para novo preenchimento.

### 1.3. Aba Paciente

Nesta aba estão disponíveis os seguintes para o preenchimento o Bloco 8 – **Informações do paciente ou usuário** e o Bloco 9 – **Evolução do Caso**.



#### 1.3.1. Bloco 8 – Informações do paciente ou usuário

Este Bloco contém os campos para o preenchimento dos dados do paciente ou usuário que sofreu o evento adverso.

8 - Informações do paciente ou usuário		
8.1. Nome completo do paciente: (*) <input type="text"/>	8.2. Iniciais: <input type="text"/>	8.3. <input type="checkbox"/> Nome e iniciais ignorados
8.4. Nome completo da mãe do paciente: <input type="text"/>	8.5. Sexo: <input type="radio"/> Masculino <input type="radio"/> Feminino	
8.6. Raça/Cor: ...: Selecione ... <input type="text"/>	8.7. Ocupação do paciente ou usuário: ...: Selecione ... <input type="text"/>	
8.8. Data de nascimento: Dia <input type="text"/> Mês <input type="text"/> Ano <input type="text"/>	8.9. Idade: <input type="text"/> ...: Selecione ... <input type="text"/>	
8.10. Número do prontuário: <input type="text"/>	8.11. Número do Cartão SUS: <input type="text"/>	
8.12. O evento levou à internação? (*) <input type="radio"/> Sim <input type="radio"/> Não <input type="radio"/> Ignorado		
8.12.1. Hospital: <input type="text"/>		
8.12.2. UF: (*) ...: Selecione ... <input type="text"/>		
8.12.3. Município: (*) ...: Selecione ... <input type="text"/>		
8.12.4. Data de Internação: Dia <input type="text"/> Mês <input type="text"/> Ano <input type="text"/>		
8.12.5. Data de Alta: Dia <input type="text"/> Mês <input type="text"/> Ano <input type="text"/>		

#### Campo 8.1. Nome completo do paciente:

Informe o nome completo do paciente ou usuário. O preenchimento deste campo gera automaticamente as informações do campo 8.2. *Iniciais*.

A identificação do paciente ou usuário será mantida sob sigilo e objetiva apenas controlar possíveis duplicidades de notificações e facilitar contato com o notificador.

#### Campo 8.2. Iniciais:

Campo preenchido automaticamente pelo sistema a partir do campo 8.1, composto das iniciais do nome completo do paciente ou usuário.

Caso o campo 8.1. não tenha sido preenchido, informe as iniciais manualmente.

#### Campo 8.3. Nome e Iniciais ignorados:

Selecione quando desconhecer as informações do nome do paciente ou usuário.

O preenchimento de pelo menos um dos campos citados acima (8.1., 8.2. ou 8.3.) é obrigatório.

#### Campo 8.4. Nome completo da mãe do paciente:

Informe o nome completo da mãe do paciente ou usuário.

A identificação da mãe do paciente ou usuário será mantida sob sigilo e objetiva apenas controlar possíveis duplicidades de notificações pela identificação de homônimos.

**Campo 8.5. Sexo:**

Selecione o sexo do paciente ou usuário que sofreu o evento adverso.

**Campo 8.6. Raça/Cor:**

Selecione uma das opções para caracterizar a raça/cor do paciente ou usuário que sofreu o evento adverso:

- Branca
- Preta
- Amarela
- Parda
- Indígena
- Ignorada

**Campo 8.7. Ocupação do paciente ou usuário:**

Selecione a ocupação do paciente ou usuário que sofreu o evento adverso.

**Campo 8.8. Data de Nascimento:**

Selecione a data de nascimento do paciente ou usuário que sofreu o evento adverso.

**Campo 8.9. Idade:**

Informe a idade do paciente ou usuário no momento da ocorrência do evento. Este campo é preenchido automaticamente pelo sistema quando a data de nascimento (campo 8.12) e a data da ocorrência (campo 4.3.) forem informadas.

**Campo 8.10. Número do prontuário:**

Informe o número do prontuário do paciente, caso exista.

**Campo 8.11. Número do Cartão SUS:**

Informe o número do Cartão SUS do paciente, caso exista.

**Campo 8.12. O evento levou à internação?**

Selecione "Sim" caso o evento tenha levado à internação. Neste caso, preencher os seguintes campos:

**Campo 8.12.1. Hospital:**

Informe o nome do hospital onde o paciente foi internado.

**Campo 8.12.2. UF:**

Selecione o estado onde se localiza o hospital em que o paciente foi internado. Esse campo deve ser selecionado antes da seleção do campo 8.12.3. *Município*.

**Campo 8.12.3. Município:**

Selecione o município onde se localiza o hospital em que o paciente foi internado. Esse campo deve ser selecionado após a seleção do campo 8.12.2. *UF*.

**Campo 8.12.4. Data de Internação:**

Selecione a data de internação.

**Campo 8.12.5. Data de Alta:**

Selecione a data de alta.

Selecione “Não” caso o evento não tenha levado à internação.

Selecione “Ignorado” caso não tenha conhecimento da ocorrência de internação em decorrência do evento.

**1.3.2. Bloco 9 – Evolução do caso**

Este bloco contém as informações sobre a evolução do caso. Essas informações serão utilizadas como subsídio para classificação de gravidade.

9 - Evolução do caso
9.1. Foi a óbito? (*) <input type="radio"/> Sim <input type="radio"/> Não
9.2. Houve lesão permanente? (*) <input type="radio"/> Sim <input type="radio"/> Não
9.3. Houve lesão temporária? (*) <input type="radio"/> Sim <input type="radio"/> Não

**Campo 9.1. Foi a óbito?**

Campo de preenchimento obrigatório.

Selecione “Sim” se o paciente evoluiu para óbito. Neste caso, a mensagem “Copiar tal qual consta na Declaração de Óbito” e os seguintes campos são apresentados:

**Campo 9.1.1. Data:**

Selecione a data em que ocorreu o óbito.

**Campo 9.1.2. Causa da morte - Campo 49 da declaração de óbito:**

Informe a causa do óbito. Este campo possibilita a inclusão de até quatro causas do óbito (itens a, b, c, d).

Selecione “Não” caso não tenha ocorrido o óbito do paciente.

### **Campo 9.2. Houve lesão permanente?**

Campo de preenchimento obrigatório.

Selecione “Sim” caso tenha ocorrido lesão permanente decorrente do evento. Neste caso, o seguinte campo é apresentado:

#### **Campo 9.2.1 Descrever:**

Descreva a lesão permanente.

Selecione “Não” caso não tenha ocorrido lesão permanente decorrente do evento.

### **Campo 9.3. Houve lesão temporária?**

Campo de preenchimento obrigatório.

Selecione “Sim” caso tenha ocorrido lesão temporária decorrente do evento. Neste caso, o seguinte campo é apresentado:

#### **Campo 9.3.1 Descrever:**

Campo alfa-numérico. Informe a descrição da lesão temporária.

Selecione “Não” caso não tenha ocorrido lesão temporária decorrente do evento.

## **1.4. Aba Outras Informações**

Nesta aba, está disponível o Bloco 10 – **Outras informações importantes**.



### **1.4.1 Bloco 10 – Outras informações importantes**

Este Bloco contém os campos para o preenchimento de informações adicionais referentes ao produto que motivou a notificação.

10 - Outras informações importantes	
10.1. A utilização do produto seguiu as instruções do fabricante?	<input type="radio"/> Sim <input type="radio"/> Não
10.2. O produto apresenta alterações (suspeita de desvio de qualidade)?	<input type="radio"/> Sim <input type="radio"/> Não <input type="radio"/> Ignorado
10.3. Local de aquisição do produto:	<input type="text" value="...: Selecione :..."/>
10.4. Possui nota fiscal da compra do produto?	<input type="radio"/> Sim <input type="radio"/> Não <input type="radio"/> Ignorado
10.5. Houve comunicação à indústria/distribuidor?	<input type="radio"/> Sim <input type="radio"/> Não <input type="radio"/> Ignorado
10.6. Foram adotadas outras providências após a identificação do problema?	<input type="radio"/> Sim <input type="radio"/> Não
10.7. Existem amostras íntegras para a coleta?	<input type="radio"/> Sim <input type="radio"/> Não
10.8. Existem rótulos do produto para a coleta?	<input type="radio"/> Sim <input type="radio"/> Não
10.9. Observações:	<input type="text"/>

**Campo 10.1. A utilização do produto seguiu as instruções do fabricante?**

Selecione a opção “Sim” caso o produto tenha sido utilizado conforme as orientações constantes no rótulo/embalagem.

Selecione a opção “Não” caso o produto não tenha sido utilizado conforme as orientações constantes no rótulo/embalagem.

**Campo 10.2. O produto apresenta alterações (suspeita de desvio da qualidade)?**

Selecione a opção “Sim” caso observe alterações no produto ou suspeite de algum desvio da qualidade. Neste caso, selecione uma ou mais opções apresentadas.



10.2. O produto apresenta alterações (suspeita de desvio de qualidade)?

Sim  Não  Ignorado

10.2.1. Quais?

- Problemas no rótulo: adulteração, ilegível, inadequado
- Presença de corpo estranho na embalagem
- Rachadura, quebra do produto ou parte dele
- Aspecto alterado: cor, manchas, cheiro
- Quantidade de unidades menor que o informado na embalagem
- Produto com data de validade vencida
- Produto com data de esterilização vencida
- Produto de uso único sendo reprocessado
- Outra

Caso a(s) alteração(ões) observada(s) não esteja(m) disponível(is) na lista, selecione a opção “Outra”. Nesta situação, o campo 10.2.1.1. Citar é apresentado para descrição.

Selecione a opção “Não” caso não observe alterações no produto ou não suspeite de desvio da qualidade.

Selecione a opção “Ignorado” caso não tenha conhecimento de alteração ou suspeita de desvio da qualidade do produto.

### Campo 10.3. Local de aquisição do produto:

Selecione o local de aquisição do produto de acordo com as opções disponíveis.

10.2. Local de aquisição do produto:

Caso o local de aquisição não esteja incluído na lista, selecione a opção “Outra”. Se a opção Outra for selecionada, preencha o campo 10.3.1 Citar para informar o local de aquisição..

10.3.1. Citar:

### Campo 10.4. Possui nota fiscal da compra do produto?

Selecione a opção “Sim” caso possua a nota fiscal de compra do produto para apresentação ao órgão do SNVS responsável, se necessário. Neste caso, será exibido em vermelho a mensagem “Mantenha-a em seu poder”.

10.4. Possui nota fiscal da compra do produto?

Sim  Não  Ignorado

Mantenha-a em seu poder.

Selecione a opção “Não” caso não possua a nota fiscal de compra do produto.

Selecione a opção “Ignorado” caso não tenha conhecimento da disponibilidade da nota fiscal de compra do produto.

#### **Campo 10.5. Houve comunicação à indústria/distribuidor?**

Selecione a opção “Sim” caso o evento adverso também tenha sido comunicado à indústria/distribuidor. Neste caso, são apresentadas as opções mostradas abaixo para seleção da forma de comunicação. Após seleção é apresentada a mensagem “Caso possua comprovante, mantenha-o em seu poder e somente entregue à Vigilância Sanitária quando solicitado.”

10.5. Houve comunicação à indústria/distribuidor?

Sim  Não  Ignorado

- Por correio
- Por e-Mail
- Por Fax
- Por representante
- Por Telefone
- Outra Forma

Caso a forma de comunicação não esteja disponível na lista, selecione a opção “Outra forma”. Nesta situação, o campo 10.5.1. Citar é apresentado para descrição.

10.5.1. Citar:

**Caso possua comprovante, mantenha-o em seu poder e somente entregue à Vigilância Sanitária quando solicitado.**

Selecione a opção “Não” caso o evento adverso não tenha sido comunicado à indústria/distribuidor.

Selecione a opção “Ignorado” caso não tenha conhecimento da comunicação do evento adverso à indústria/distribuidor.

#### **Campo 10.6. Foram adotadas outras providências após a identificação do problema?**

Selecione a opção “Sim” caso tenham sido adotadas outras providências além das descritas neste formulário de notificação. Exemplo: suspensão do uso e separação do lote. Nesta situação, preencher o campo 10.6.1 Citar para informar as outras medidas adotadas.

10.6.1. Citar:

Selecione a opção “Não” caso não tenham sido adotadas outras providências, além das descritas neste formulário de notificação.

#### **Campo 10.7. Existem amostras íntegras para a coleta?**

Selecione a opção “Sim” caso possua amostras íntegras do produto disponíveis para coleta pelo órgão do SNVS responsável, se necessário. Neste caso, o campo 10.7.1 *Quantas?* e a mensagem “Mantenha-as em seu poder e somente as entregue à Vigilância Sanitária quando solicitado” são apresentados.

10.7.1. Quantas?

Mantenha-as em seu poder e somente as entregue à Vigilância Sanitária quando solicitado.

Selecione a opção “Não” caso não possua amostras íntegras do produto disponíveis para coleta.

### Campo 10.8. Existem rótulos do produto para a coleta?

Selecione a opção “Sim” caso possua rótulos do produto disponíveis para coleta pelo órgão do SNVS responsável, se necessário. Neste caso, a mensagem “Mantenha-os em seu poder e somente os entregue à Vigilância Sanitária quando solicitado” é apresentada.

Selecione a opção “Não” caso não possua rótulos do produto disponíveis para coleta.

### Campo 10.9. Observações:

Caso considere relevante alguma informação ainda não relatada, informe neste campo.

## 1.5. Aba Pendências

A qualquer momento as pendências poderão ser verificadas selecionando esta aba.

Nesta Aba, são apontadas as pendências quanto à violação de regras e obrigações no preenchimento do formulário de notificação. Está disponível o Bloco **Pendências**.

Motivo	Produto e Empresa	Paciente	Outras Informações	Pendências
--------	-------------------	----------	--------------------	------------

### 1.5.1. Bloco – Pendências

Este Bloco apresenta os campos que necessitam de verificação para o notificador providenciar a correção ou preenchimento antes do envio.

Pendências
Os seguintes campos necessitam verificação:
<b>Motivo:</b>
Data da identificação: Campo obrigatório não preenchido.
Descreva detalhadamente a Queixa Técnica/Evento Adverso: Campo obrigatório não preenchido.
Município: Campo obrigatório não preenchido.
Selecione uma das opções: Campo obrigatório não preenchido.

Após preenchimento do formulário, caso não haja pendências, será apresentada a mensagem “A Notificação foi verificada e não existe pendência”.

Para enviar a notificação, clique no botão  . O certificado de envio será apresentado:

**Notificação Enviada**

**Número da Notificação:** 2006.10.001808  
**Data da notificação:** 24/10/2006  
**Notificador:** bruno TécnicoSNVS  
**Produto motivo da notificação:** Artigo

Caro(a) Notificador(a),  
sua notificação foi recebida com sucesso e será analisada pelos técnicos da Vigilância Sanitária.  
Atenciosamente,

NOTIVISA - Sistema Nacional de Notificações para a Vigilância Sanitária

Para imprimir o formulário preenchido, clique em  e para imprimir o certificado, clique em .

## Anexo II

# Formulário para Notificação de Queixa Técnica de Artigo Médico-Hospitalar

# Manual do Usuário



Notivisa – Sistema Nacional de Notificações  
para a Vigilância Sanitária

## Anexo 7

**Formulário para Notificação de Queixa Técnica de Artigo  
Médico-Hospitalar**

## Sumário

1. Formulário de Notificação de Queixa Técnica de Artigo Médico-Hospitalar. . .	3
<b>1.1. Aba Motivo</b> .....	<b>3</b>
1.1.1. Bloco 3 – Tipo de Queixa Técnica .....	3
1.1.2. Bloco 4 – Queixa Técnica .....	4
<b>1.2. Aba Produto e Empresa</b> .....	<b>6</b>
1.2.1. Bloco 5 – Produto e empresa .....	6
1.2.2. Bloco 6– Dados do Produto .....	7
1.2.3. Bloco 7– Dados do Fabricante (seleção da opção “Não” no campo 6.9).....	8
1.2.4. Bloco 7– Dados do importador (seleção da opção “Sim” no campo 6.9).....	9
<b>1.3. Aba Outras Informações</b> .....	<b>11</b>
1.3.1. Bloco 10 – Outras informações importantes.....	11
<b>1.4. Aba Pendências</b> .....	<b>13</b>
1.4.1. Bloco – Pendências .....	14

## 1. Formulário de Notificação de Queixa Técnica de Artigo Médico-Hospitalar

Para Queixa Técnica de Artigo Médico-Hospitalar (implante ortopédico, DIU, válvula cardíaca, seringa, catéter, equipo, etc.) estão disponíveis as seguintes abas:

- Motivo;
- Produto e Empresa;
- Outras informações;
- Pendências.

### 1.1. Aba Motivo

Formulário para notificação de Queixa Técnica de Artigo Médico-Hospitalar  
Número da Notificação: 2006.12.000030



Motivo | Produto e Empresa | Outras Informações | Pendências

#### 1.1.1. Bloco 3 – Tipo de Queixa Técnica

Este Bloco contém as opções para caracterização da queixa técnica a ser notificada.

**3 - Tipo de Queixa Técnica**

3.1. Selecione uma das opções: (\*)

...: Selecione ...

Selecione uma das opções disponíveis no campo 3.1. Este campo é de preenchimento obrigatório.

Opção	Descrição
Produto com suspeita de desvio da qualidade	Selecione quando houver suspeita de afastamento dos parâmetros de qualidade estabelecidos para um produto ou processo.
Produto com suspeita de estar sem Registro	Selecione quando houver suspeita de que o produto não esteja registrado ou tenha o registro vencido. O número de registro no Ministério da Saúde (MS) deve estar impresso no rótulo/embalagem do produto, após a sigla MS.
Suspeita de empresa sem autorização de funcionamento (AFE)	Selecione quando houver suspeita de que a empresa fabricante ou importadora do produto esteja sem autorização de funcionamento (AFE), que é emitida pelo órgão ou entidade competente do Ministério da Saúde.
Suspeita de produto falsificado	Selecione quando houver suspeita de que o produto tenha sido fabricado e comercializado sem o conhecimento ou consentimento da Anvisa/MS e que simula a identidade de outro produto já legalmente registrado.
Suspeita de outras práticas irregulares	Selecione para indicar suspeita de outras práticas não citadas nas opções acima.



Quando a opção “Produto com suspeita de desvio de qualidade” for selecionada, assinale uma das opções do campo 3.1.1. *O produto apresenta alterações?*, conforme segue:

3 - Tipo de Queixa Técnica	
3.1. Selecione uma das opções: (*)	
Produto com suspeita de desvio da qualidade	
3.1.1. O produto apresenta alterações?	
<input type="checkbox"/>	Problemas no rótulo: adulteração, ilegível, inadequado
<input type="checkbox"/>	Presença de corpo estranho na embalagem
<input type="checkbox"/>	Rachadura, quebra do produto ou parte dele
<input type="checkbox"/>	Aspecto alterado: cor, manchas, cheiro
<input type="checkbox"/>	Quantidade de unidades menor que o informado na embalagem
<input type="checkbox"/>	Produto com data de validade vencida
<input type="checkbox"/>	Produto com data de esterilização vencida
<input type="checkbox"/>	Produto de uso único sendo reprocessado
<input type="checkbox"/>	Outra

Caso a alteração não esteja disponível na lista, selecione a opção “Outra”. Nesta situação, o campo 3.1.2. *Citar:* é apresentado para descrição.

3.1.2. Citar:

Quando a opção “Suspeita de outras práticas irregulares” for assinalada, o campo 3.1.2. *Citar:* é apresentado para descrição.

3.1. Selecione uma das opções: (\*)

Suspeita de outras práticas irregulares

3.1.2. Citar:

### 1.1.2. Bloco 4 – Queixa Técnica

Este Bloco contém campos para preenchimento de dados de identificação e descrição da Queixa Técnica.

4 - Queixa Técnica		
4.1. Descreva detalhadamente a Queixa Técnica : (*)		
4.3. Data da identificação: (*)		
Dia	Mês	Ano
4.4. Local da identificação:		4.5. Endereço do local da identificação:
Estabelecimento de saúde		515N - Edifício ANVISA
4.4.1. Nome do estabelecimento de saúde:		
Empresa de Teste Ltda		
4.4.2. CNPJ do estabelecimento de saúde:		
33.683.202/0001-34		
4.6. País: (*)	4.7. UF: (*)	4.8. Município: (*)
BRASIL	Distrito Federal - [	BRASÍLIA

**Campo 4.1. Descreva detalhadamente a Queixa Técnica:**

Campo de preenchimento obrigatório. Descreva detalhadamente a queixa técnica incluindo os motivos da notificação.

**Campo 4.3. Data da identificação:**

Campo de preenchimento obrigatório. Selecione a data da identificação do problema (queixa técnica).

**Campo 4.4. Local da identificação:**

Selecione o local onde o problema foi identificado dentre as seguintes opções:

- Estabelecimento de Saúde;
- Residência;
- Outros.

Caso selecione a opção “Estabelecimento de Saúde”, preencha os seguintes campos:

- 4.4.1. Nome do estabelecimento de saúde:
- 4.4.2. CNPJ do estabelecimento de saúde:

Caso selecione a opção “Outro”, preencha o seguinte campo:

- 4.4.3. Citar:

**Campo 4.5. Endereço do local da identificação:**

Campo alfa-numérico. Informe o endereço do local da identificação do problema (queixa técnica).

**Campo 4.6. País:**

Campo de preenchimento obrigatório. Selecione o país onde ocorreu o problema (queixa técnica). A opção “Brasil” está pré-selecionada.

### Campo 4.7. UF:

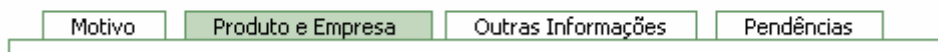
Selecione o estado onde foi identificado o problema. Este campo deve ser selecionado antes da seleção do campo 4.8. *Município*. Se o país selecionado no campo 4.6 for “Brasil”, este campo é de preenchimento obrigatório.

### Campo 4.8. Município:

Selecione o município onde foi identificado o problema. Só é possível selecionar este campo após a seleção do campo 4.7. *UF*. Se o país selecionado no campo 4.6 for “Brasil”, este campo é de preenchimento obrigatório.

## 1.2. Aba Produto e Empresa

Nesta Aba, estão disponíveis os seguintes Blocos: **Produto e Empresa**, **Dados do Produto** e **Dados do Fabricante ou Importador**.



### 1.2.1. Bloco 5 – Produto e empresa

Este Bloco contém os campos para preenchimento do Número do Registro do produto ou cadastro na ANVISA/MS e do CNPJ da empresa fabricante ou importadora. Para facilitar o preenchimento de dados desta aba, opcionalmente, o notificador poderá realizar a pesquisa no banco de dados da Anvisa.

5 - Produto e Empresa	
5.1. Número do registro ou cadastro na ANVISA/MS: <input type="text"/>	5.2. CNPJ da empresa fabricante ou importador: <input type="text"/>
<input type="button" value="Pesquisar"/>	<input type="button" value="Pesquisar"/>

Dependendo da pesquisa escolhida, os campos do Bloco 5 – **Produto e Empresa**, Bloco 6 - **Dados do produto** e Bloco 7 – **Dados do Fabricante ou Importador** são preenchidos com os dados resultantes da pesquisa.

Estes Campos podem ser preenchidos automaticamente por pesquisa realizada em campos do Bloco 6 – **Dados do Produto**.

#### Campo 5.1. Número do registro ou cadastro na ANVISA/MS:

Campo numérico com espaço para até 13 dígitos. Informe o número do registro do produto na ANVISA/MS, impresso no rótulo ou embalagem do produto.

Após preenchimento, clique no botão “pesquisar”.

Caso os dados do produto sejam encontrados no banco de dados da Anvisa, alguns campos referentes ao produto e empresa serão preenchidos automaticamente.

Estes dados devem ser verificados e, caso não confirmem com o produto em mãos, devem ser alterados manualmente.

### Campo 5.2. CNPJ da empresa fabricante ou importador:

Campo numérico com espaço para até 14 dígitos. Informe o CNPJ da empresa fabricante ou importadora, impresso no rótulo/embalagem do produto.

Após preenchimento, clique no botão “pesquisar”.

Caso os dados da empresa sejam encontrados no banco de dados da Anvisa, alguns campos referentes ao produto e empresa serão preenchidos automaticamente. Estes dados devem ser verificados e, caso não confirmem com o produto em mãos, devem ser alterados manualmente.

### 1.2.2. Bloco 6– Dados do Produto

Este Bloco contém os campos para o preenchimento dos dados referentes ao produto, que poderão vir preenchidos automaticamente como resultado da pesquisa selecionada no Bloco 5.

6 - Dados do Produto	
6.1. Nome comercial do produto:	
<input type="text"/>	<input type="button" value="Pesquisar"/>
6.2. Produto:	
...Selecione:...	
6.3. Data de fabricação:	6.4. Data de validade:
<input type="text"/> Dia <input type="text"/> Mês <input type="text"/> Ano	<input type="text"/> Dia <input type="text"/> Mês <input type="text"/> Ano
6.5. Data de esterilização:	6.6. Data de validade da esterilização:
<input type="text"/> Dia <input type="text"/> Mês <input type="text"/> Ano	<input type="text"/> Dia <input type="text"/> Mês <input type="text"/> Ano
6.7. Número do lote/série:	6.8. O produto foi reprocessado?
<input type="text"/>	<input type="radio"/> Sim <input type="radio"/> Não
6.9. Produto importado?	
<input type="radio"/> Sim <input type="radio"/> Não	

### Campo 6.1. Nome comercial do produto:

Campo para identificação do produto. Preencha com o nome comercial do produto constante na rótulo/embalagem do produto.

Para a realização da pesquisa a partir deste campo, pelo menos 3 (três) caracteres devem ser informados.

Após clicar em “Pesquisar”, selecione uma das opções da lista apresentada. Caso os dados confirmem com o produto em mãos, clique no botão  e preencha todos os campos manualmente.

Este campo pode ser obrigatório, dependendo da seleção do tipo de queixa técnica, no campo 3.1, da aba Motivo.

Caso tenha sido realizada pesquisa pelo CNPJ, no campo 5.2, este campo poderá apresentar lista com produtos fabricados pela empresa em questão para seleção do produto que motivou a notificação.

### Campo 6.2. Produto:

Campo para seleção da categoria do produto. Este campo não pode ser preenchido manualmente.

Após a realização de pesquisa nos campos 5.1., 5.2. ou 6.1., o sistema poderá carregar automaticamente uma tabela para a seleção da categoria do produto. Caso não haja opções para seleção, desconsidere este campo.

**Campo 6.3. Data de fabricação:**

Informe a data de fabricação do produto, que consta em seu rótulo/embalagem.

**Campo 6.4. Data de validade:**

Informe a data de validade do produto, que consta em seu rótulo/embalagem.

**Campo 6.5. Data de esterilização:**

Informe a data de esterilização do produto.

**Campo 6.6. Data de validade da esterilização:**

Informe a data de validade da esterilização.

**Campo 6.7. Número do lote/série:**

Campo alfa-numérico com espaço para até 14 dígitos. Informe o número do lote/série do produto.

**Campo 6.8. O produto foi reprocessado?**

Selecione a opção “Sim” caso o produto tenha sido reprocessado. Nesse caso preencha o seguinte campo:

**Campo 6.8.1. “Quantas vezes foi reprocessado antes do incidente?”**

Selecione a opção “Não” caso o produto não tenha sido reprocessado.

**Campo 6.9. Produto importado?**

Selecione a opção “Sim” caso o produto seja importado. Esta informação pode ser verificada no rótulo/embalagem do produto.

Selecione a opção “Não” caso o produto seja nacional. Esta informação pode ser verificada no rótulo/embalagem do produto.

**1.2.3. Bloco 7– Dados do Fabricante (seleção da opção “Não” no campo 6.9)**

Este Bloco contém os campos para preenchimento dos dados do fabricante.

Caso tenha sido realizada pesquisa nos campos anteriores, os campos deste bloco poderão ser preenchidos automaticamente.

6.9. Produto importado?	
<input type="radio"/> Sim <input checked="" type="radio"/> Não	
<b>7 - Dados do fabricante</b>	
7.1. Nome ou razão social do fabricante:	
<input type="text"/>	
7.2. Endereço do fabricante:	
<input type="text"/>	
7.3. Número do telefone/SAC do fabricante:	
<input type="text"/>	
7.4. UF do fabricante:	7.5. Município do fabricante:
<input type="text" value="...: Selecione ..."/>	<input type="text" value="...: Selecione ..."/>
<input type="button" value="Limpar dados da empresa e do produto"/>	

**Campo 7.1. Nome ou razão social do fabricante:**

Informe o nome ou sua razão social do fabricante que consta no rótulo/embalagem do produto.

Este campo pode ser obrigatório, dependendo da seleção do tipo de queixa técnica, no campo 3.1, da aba Motivo.

**Campo 7.2. Endereço do fabricante:**

Informe o endereço do fabricante que consta no rótulo/embalagem do produto.

**Campo 7.3. Número do telefone/SAC do fabricante:**

Informe o número do telefone ou Serviço de Atendimento ao Cliente (SAC) do fabricante que consta no rótulo/embalagem do produto.

**Campo 7.4. U.F. do fabricante:**

Selecione o estado do fabricante que consta no rótulo/embalagem do produto. Este campo deve ser selecionado antes da seleção do campo 7.5. *Município do fabricante.*

**Campo 7.5. Município do fabricante:**

Selecione o município do fabricante que consta no rótulo/embalagem do produto. Apenas é possível selecionar este campo após a seleção do campo 7.4. *UF do fabricante.*

**1.2.4. Bloco 7– Dados do importador (seleção da opção “Sim” no campo 6.9)**

Este Bloco contém os campos para preenchimento dos dados do importador.

6.9. Produto importado? <input checked="" type="radio"/> Sim <input type="radio"/> Não	
<b>7 - Dados do importador</b>	
7.1. Nome ou razão social do importador: <input type="text"/>	
7.2. Endereço do importador: <input type="text"/>	
7.3. Número do telefone/SAC do importador: <input type="text"/>	
7.4. UF do importador: <input type="text" value="...: Selecione ..."/>	7.5. Município do importador: <input type="text" value="...: Selecione ..."/>
7.6. Nome do fabricante: <input type="text"/>	7.7. País do fabricante: <input type="text" value="...: Selecione ..."/>
<input type="button" value="Limpar dados da empresa e do produto"/>	

**Campo 7.1. Nome ou razão social do importador:**

Informe o nome ou sua razão social do importador que consta no rótulo/embalagem do produto.

Este campo pode ser obrigatório, dependendo da seleção do tipo de queixa técnica, no campo 3.1, da aba Motivo.

**Campo 7.2. Endereço do importador:**

Informe o endereço do importador que consta no rótulo/embalagem do produto.

**Campo 7.3. Número do telefone/SAC do importador:**

Informe com o número do telefone ou Serviço de Atendimento ao Cliente (SAC) do importador que consta no rótulo/embalagem do produto.

**Campo 7.4. U.F. do importador:**

Selecione o estado do importador que consta no rótulo/embalagem do produto. Este campo deve ser selecionado antes da seleção do campo 7.5. *Município do importador.*

**Campo 7.5. Município do importador:**

Selecione o município do importador que consta no rótulo/embalagem do produto. Apenas é possível selecionar este campo após a seleção do campo 7.4. *UF do importador.*

**Campo 7.6. Nome do fabricante:**

Informe o nome ou razão social do fabricante que consta no rótulo/embalagem do produto.

**Campo 7.7. País do Fabricante:**

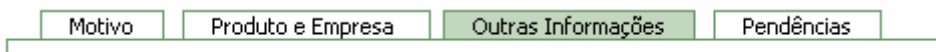
Selecione o país do fabricante do produto que consta no rótulo/embalagem.

**Botão Limpar dados da empresa e do produto**

Caso haja inconsistência ou incoerência no resultado da pesquisa realizada, clique nesse botão e os dados preenchidos nesta aba serão limpos para novo preenchimento.

### 1.3. Aba Outras Informações

Nesta aba, está disponível o **Bloco 10 – Outras informações importantes**.



#### 1.3.1. Bloco 10 – Outras informações importantes

Este Bloco contém os campos para o preenchimento de informações adicionais referentes ao produto que motivou a notificação.

10 - Outras informações importantes	
10.1. A utilização do produto seguiu as instruções do fabricante?	<input type="radio"/> Sim <input type="radio"/> Não
10.3. Local de aquisição do produto:	...: Selecione ...
10.4. Possui nota fiscal da compra do produto?	<input type="radio"/> Sim <input type="radio"/> Não <input type="radio"/> Ignorado
10.5. Houve comunicação à indústria/distribuidor?	<input type="radio"/> Sim <input type="radio"/> Não <input type="radio"/> Ignorado
10.6. Foram adotadas outras providências após a identificação do problema?	<input type="radio"/> Sim <input type="radio"/> Não
10.7. Existem amostras íntegras para a coleta?	<input type="radio"/> Sim <input type="radio"/> Não
10.8. Existem rótulos do produto para a coleta?	<input type="radio"/> Sim <input type="radio"/> Não
10.9. Observações:	<div style="border: 1px solid black; height: 40px;"></div>

#### Campo 10.1. A utilização do produto seguiu as instruções do fabricante?

Selecione a opção “Sim” caso o produto tenha sido utilizado conforme informações constantes, no rótulo/embalagem.

Selecione a opção “Não” caso o produto não tenha sido utilizado conforme informações constantes no rótulo/embalagem. Esta opção também deve ser selecionada quando foi utilizado um procedimento rotineiro e bem conhecido não descrito pelo fabricante.

#### Campo 10.3. Local de aquisição do produto:

Selecione o local de aquisição do produto de acordo com as opções disponíveis.




10.3. Local de aquisição do produto:



Caso o local de aquisição não esteja disponível na lista, selecione a opção "Outro". Nesta situação, o campo 10.3.1 Citar é apresentado para descrição.

10.3. Local de aquisição do produto:



#### Campo 10.4. Possui nota fiscal da compra do produto?

Selecione a opção "Sim" caso possua a nota fiscal de compra do produto disponível para apresentação ao órgão do SNVS responsável, se necessário. Neste caso, será apresentada a mensagem "Mantenha-a em seu poder" em vermelho.

10.4. Possui nota fiscal da compra do produto?

Sim
  Não
  Ignorado

**Mantenha-a em seu poder.**

Selecione a opção "Não" caso não possua a nota fiscal de compra do produto

Selecione a opção "Ignorado" caso não tenha conhecimento da disponibilidade da nota fiscal de compra do produto.

#### Campo 10.5. Houve comunicação à indústria/distribuidor?

Selecione a opção "Sim" caso a queixa técnica também tenha sido informada à indústria/distribuidor. Neste caso, são apresentadas as opções mostradas abaixo para seleção da forma de comunicação. Após seleção é apresentada a mensagem "Caso possua comprovante, mantenha-o em seu poder e somente entregue à Vigilância Sanitária quando solicitado".

10.5. Houve comunicação à indústria/distribuidor?

Sim
  Não
  Ignorado

Por correio

Por e-Mail

Por Fax

Por representante

Por Telefone

Outra Forma

**Caso possua comprovante, mantenha-o em seu poder e somente entregue à Vigilância Sanitária quando solicitado.**

Caso a forma de comunicação não esteja disponível na lista, selecione a opção "Outra forma". Nesta situação, o campo 10.5.1 Citar é apresentado para descrição.

10.5.1. Citar:

Selecione a opção “Não” caso a queixa técnica não tenha sido comunicada à indústria/distribuidor.

Selecione a opção “Ignorado” caso não tenha conhecimento da comunicação da queixa técnica à indústria/distribuidor.

**Campo 10.6. Foram adotadas outras providências após a identificação do problema?**

Selecione a opção “Sim” caso tenham sido adotadas outras providências, além das descritas neste formulário de notificação. Nesta situação, o campo *10.6.1 Citar* é apresentado para descrição.

10.6.1. Citar:

Selecione a opção “Não” caso não tenham sido adotadas outras providências, além das descritas neste formulário de notificação.

**Campo 10.7. Existem amostras íntegras para a coleta?**

Selecione a opção “Sim” caso possua amostras íntegras do produto disponíveis para coleta pelo órgão do SNVS responsável, se necessário. Neste caso, o campo *10.7.1 Quantas?* e a mensagem “Mantenha-as em seu poder e somente as entregue à Vigilância Sanitária quando solicitado” são apresentados.

10.7.1. Quantas?

Mantenha-as em seu poder e somente as entregue à Vigilância Sanitária quando solicitado.

Selecione a opção “Não” caso não possua amostras íntegras do produto disponíveis para coleta.

**Campo 10.8. Existem rótulos do produto para a coleta?**

Selecione a opção “Sim” caso possua rótulos do produto disponíveis para coleta pelo órgão do SNVS responsável, se necessário. Neste caso, a mensagem “Mantenha-os em seu poder e somente os entregue à Vigilância Sanitária quando solicitado” é apresentada.

Selecione a opção “Não” caso não possua rótulos do produto disponíveis para coleta.

**Campo 10.9. Observações:**

Caso considere relevante alguma informação ainda não relatada, informe neste campo.

**1.4. Aba Pendências**

A qualquer momento as pendências poderão ser verificadas selecionando esta aba.

Nesta Aba, são apontadas as pendências quanto à violação de regras e obrigatoriedades no preenchimento do formulário de notificação. Dispõe de apenas um Bloco **Pendências**.

Motivo	Produto e Empresa	Outras Informações	Pendências
--------	-------------------	--------------------	------------

### 1.4.1. Bloco – Pendências

Este Bloco apresenta os campos que necessitam de verificação para o notificador providenciar a correção ou preenchimento antes do envio.

Pendências
Os seguintes campos necessitam verificação: <b>Motivo:</b> Data da identificação: Campo obrigatório não preenchido. Descreva detalhadamente a Queixa Técnica/Evento Adverso: Campo obrigatório não preenchido. Município: Campo obrigatório não preenchido. Selecione uma das opções: Campo obrigatório não preenchido.

Após preenchimento do formulário, caso não haja pendências, será apresentada a mensagem “A Notificação foi verificada e não existe pendência”.

Para enviar a notificação, clique no botão  . O certificado de envio será apresentado:

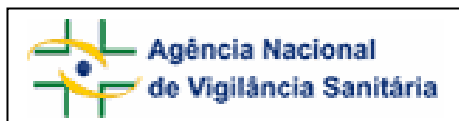
Notificação Enviada
<b>Número da Notificação:</b> 2006.10.001808 <b>Data da notificação:</b> 24/10/2006 <b>Notificador:</b> bruno Tecnico5NVS <b>Produto motivo da notificação:</b> Artigo
Caro(a) Notificador(a), sua notificação foi recebida com sucesso e será analisada pelos técnicos da Vigilância Sanitária. Atenciosamente,
NOTIVISA - Sistema Nacional de Notificações para a Vigilância Sanitária
<input type="button" value="Imprimir Notificação"/> <input type="button" value="Imprimir Certificado"/>

Para imprimir o formulário preenchido, clique em  e para imprimir o certificado, clique em  .

## **Anexo III**

# **Formulário para Notificação de Evento Adverso Associado ao uso de Equipamento Médico-hospitalar**

# Manual do Usuário



Notivisa – Sistema Nacional de Notificações  
para a Vigilância Sanitária

## Anexo 10

**Formulário para Notificação de Evento Adverso Associado ao  
uso de Equipamento Médico-hospitalar**

## Sumário

<b>1. Formulário para Notificação de Evento Adverso associado ao uso de Equipamento Médico-Hospitalar .....</b>	<b>3</b>
<b>1.1. ABA MOTIVO.....</b>	<b>3</b>
1.1.1. <i>Bloco 4 – Evento Adverso.....</i>	<i>3</i>
<b>1.2. ABA PRODUTO E EMPRESA .....</b>	<b>5</b>
1.2.1. <i>Bloco 5 – Produto e empresa .....</i>	<i>5</i>
1.2.2. <i>Bloco 6– Dados do Produto.....</i>	<i>6</i>
1.2.3. <i>Bloco 7– Dados do Fabricante (seleção da opção “Não”) .....</i>	<i>7</i>
1.2.4. <i>Bloco 7– Dados do Fabricante (seleção da opção “Sim”) .....</i>	<i>8</i>
<b>1.3. ABA PACIENTE .....</b>	<b>10</b>
1.3.1. <i>Bloco 8 – Informações do paciente ou usuário.....</i>	<i>10</i>
1.3.2. <i>Bloco 9 – Evolução do caso.....</i>	<i>12</i>
<b>1.4. ABA OUTRAS INFORMAÇÕES.....</b>	<b>13</b>
1.4.1. <i>Bloco 10 – Outras informações importantes .....</i>	<i>13</i>
<b>1.5. ABA PENDÊNCIAS.....</b>	<b>17</b>
1.5.1. <i>Bloco – Pendências.....</i>	<i>17</i>

## 1. Formulário para Notificação de Evento Adverso associado ao uso de Equipamento Médico-Hospitalar

Para Evento Adverso de Equipamento Médico-Hospitalar (Marcapasso, Ventilador Pulmonar, Bomba de Infusão, etc.) estão disponíveis as seguintes abas:

- Motivo;
- Produto e Empresa;
- Paciente;
- Outras informações;
- Pendências.

### 1.1. Aba Motivo

Formulário para notificação de Evento Adverso associado ao uso de Equipamento Médico-Hospitalar  
Número da Notificação: 2006.11.000066

Motivo

Produto e Empresa

Paciente

Outras Informações

Pendências

#### 1.1.1. Bloco 4 – Evento Adverso

Este Bloco contém os campos para descrição e detalhamento do Evento Adverso.

4 - Evento Adverso		
4.1. Descreva detalhadamente o Evento Adverso : (*)		
<input type="text"/>		
4.2. Seleccione o evento adverso (WHO-ART):		
<input type="text"/>		<input type="button" value="Pesquisar"/>
4.3. Data da ocorrência: (*)		
<input type="text" value="Dia"/>	<input type="text" value="Mês"/>	<input type="text" value="Ano"/>
4.4. Local da ocorrência:		4.5. Endereço do local da ocorrência:
<input type="text" value="Estabelecimento de saúde"/>		<input type="text" value="515N - Edifício ANVISA"/>
4.4.1. Nome do estabelecimento de saúde:		
<input type="text" value="Empresa de Teste Ltda"/>		
4.4.2. CNPJ do estabelecimento de saúde:		
<input type="text" value="33.683.202/0001-34"/>		
4.6. País: (*)	4.7. UF: (*)	4.8. Município: (*)
<input type="text" value="BRASIL"/>	<input type="text" value="...: Seleccione ..."/>	<input type="text" value="...: Seleccione ..."/>


#### Campo 4.1. Descreva detalhadamente o Evento Adverso:

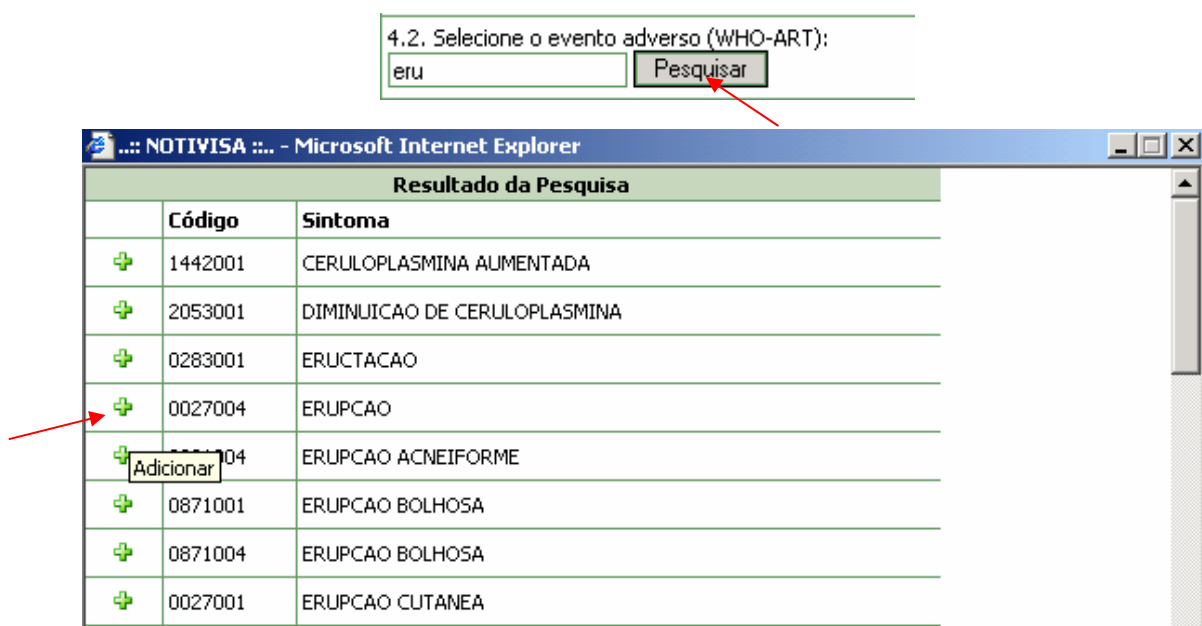
Campo de preenchimento obrigatório. Descreva detalhadamente o evento adverso incluindo os motivos da notificação.

#### Campo 4.2. Selecione o evento adverso (WHO-ART):

Campo para pesquisa da terminologia padronizada de evento adverso no WHO-ART (World Health Organization – Adverse Reaction Terminology).

Para a realização desta pesquisa digite, pelo menos, 3 caracteres do sinal/sintoma a ser pesquisado e clique em “pesquisar”.

Após aberta a tela *pop-up* com as opções, selecione a opção desejada clicando no símbolo , localizado à esquerda da tela. Veja o exemplo:



Caso queira selecionar outro sinal/sintoma, repita a operação.

#### Campo 4.3. Data da ocorrência:

Campo de preenchimento obrigatório. Informe a data da ocorrência do evento adverso.

#### Campo 4.4. Local da ocorrência:

Selecione o local da ocorrência do evento adverso dentre as seguintes opções:

- Estabelecimento de Saúde;
- Residência;
- Outros.

Caso selecione a opção “Estabelecimento de Saúde”, preencha os seguintes campos:

- 4.4.1. Nome do estabelecimento de saúde:
- 4.4.2. CNPJ do estabelecimento de saúde:

Caso selecione a opção “Outro”, preencha o seguinte campo:



4.4.3. Citar:

**Campo 4.5. Endereço do local da ocorrência:**

Informe o endereço do local da ocorrência do evento adverso.

**Campo 4.6. País:**

Campo de preenchimento obrigatório. Informe o país onde ocorreu o evento adverso. A opção “Brasil” está pré-selecionada.

**Campo 4.7. UF:**

Informe o estado onde ocorreu o evento adverso. Este campo deve ser selecionado antes da seleção do campo 4.8. *Município*. Se o país selecionado no campo 4.6 for “Brasil”, este campo é de preenchimento obrigatório.

**Campo 4.8. Município:**

Informe o município onde ocorreu o evento adverso. É possível selecionar este campo somente após a seleção do campo 4.7. *UF*. Se o país selecionado no campo 4.6 for “Brasil”, este campo é de preenchimento obrigatório.

## 1.2. Aba Produto e Empresa

Nesta Aba, estão disponíveis os seguintes Blocos: “**Produto e Empresa**”, “**Dados do Produto**” e “**Dados do Fabricante ou Importador**”.



### 1.2.1. Bloco 5 – Produto e empresa

Este Bloco contém os campos para preenchimento do número do registro ou cadastro na ANVISA/MS e do CNPJ da empresa fabricante ou importadora. Para facilitar o preenchimento de dados desta aba, opcionalmente, o notificador poderá realizar a pesquisa no banco de documentos da Anvisa.

5 - Produto e Empresa	
5.1. Número do registro ou cadastro na ANVISA/MS: <input type="text"/>	5.2. CNPJ da empresa fabricante ou importador: <input type="text"/>
<input type="button" value="Pesquisar"/>	<input type="button" value="Pesquisar"/>

Dependendo da pesquisa realizada, os campos do Bloco 5 – **Produto e Empresa**, Bloco 6 - **Dados do produto** e Bloco 7 – **Dados do Fabricante ou Importador** poderão ser preenchidos automaticamente.

Estes campos podem ser preenchidos automaticamente por pesquisa realizada em campos do Bloco 6 – **Dados do Produto**.

**Campo 5.1. Número do registro na ANVISA/MS:**

Campo numérico com espaço para até 13 dígitos. Informe o número do registro do equipamento na ANVISA/MS, impresso no rótulo/embalagem ou no próprio equipamento.

Após preenchimento, clique no botão “Pesquisar”.

Caso os dados do equipamento sejam encontrados no banco de dados da Anvisa, alguns campos referentes ao equipamento e empresa são preenchidos automaticamente. Estes dados devem ser verificados e, caso não confirmem com o produto, devem ser alterados manualmente.

Estes dados devem ser verificados e, caso não confirmem com o produto associado à notificação do evento adverso, devem ser alterados manualmente pelo notificador.

### Campo 5.2. CNPJ da empresa fabricante ou importador:

Campo numérico com espaço para até 14 dígitos. Informe o CNPJ da empresa fabricante ou importadora impresso no equipamento.

Após preenchimento, clique no botão “pesquisar”.

Caso os dados da empresa sejam encontrados no banco de dados da Anvisa, alguns campos referentes ao produto e empresa são preenchidos automaticamente. Estes dados devem ser verificados e, caso não confirmem com o produto, devem ser alterados manualmente.

## 1.2.2. Bloco 6– Dados do Produto

Este Bloco contém os campos para o preenchimento dos dados referentes ao produto, que poderão ser preenchidos automaticamente como resultado da pesquisa realizada no Bloco 5.

5 - Produto e Empresa	
5.1. Número do registro ou cadastro na ANVISA/MS: <input type="text"/>	5.2. CNPJ da empresa fabricante ou importador: <input type="text"/>
<input type="button" value="Pesquisar"/>	<input type="button" value="Pesquisar"/>
6 - Dados do Produto	
6.1. Nome comercial do equipamento: <input type="text"/>	
<input type="button" value="Pesquisar"/>	
6.2. Equipamento: [...:Selecione:...] <input type="button" value="v"/>	
6.3. Modelo do equipamento: <input type="text"/>	6.4. Número de série de fabricação: <input type="text"/>
6.5. Data de aquisição: Dia <input type="text"/> Mês <input type="text"/> Ano <input type="text"/>	6.6. Data de instalação: Dia <input type="text"/> Mês <input type="text"/> Ano <input type="text"/>
6.7. Data do término da garantia: Dia <input type="text"/> Mês <input type="text"/> Ano <input type="text"/>	
6.8. Equipamento importado? <input type="radio"/> Sim <input type="radio"/> Não	

### Campo 6.1. Nome comercial do equipamento:

Campo para identificação do equipamento. Preencha com o nome comercial do equipamento constante na sua embalagem ou manual de instrução.

Para a realização da pesquisa a partir deste campo, pelo menos 3 (três) caracteres devem ser informados.

Após clicar em “Pesquisar”, selecione uma das opções da lista apresentada. Caso os dados confirmem com o produto em mãos, clique no botão  e preencha todos os campos manualmente.

Este campo pode ser obrigatório, dependendo da seleção do tipo de queixa técnica, no campo 3.1, da aba Motivo.

Caso tenha sido realizada pesquisa pelo CNPJ, no campo 5.2, este campo poderá apresentar lista com equipamentos fabricados pela empresa em questão para seleção do equipamento que motivou a notificação.

**Campo 6.2. Equipamento:**

Campo para seleção da categoria do equipamento. Este campo não pode ser preenchido manualmente. Após a realização de pesquisa nos campos 5.1., 5.2. ou 6.1., o sistema poderá carregar automaticamente uma tabela para a seleção da categoria do equipamento. Caso não haja opções para seleção, desconsidere este campo.

**Campo 6.3. Modelo do Equipamento:**

Informe o modelo do equipamento.

**Campo 6.4. Número de série de fabricação:**

Informe o número de série de fabricação.

**Campo 6.5. Data de aquisição:**

Informe a data de aquisição do equipamento.

**Campo 6.6. Data de instalação:**

Informe a data de instalação de instalação do equipamento no serviço de saúde.

**Campo 6.7. Data do Término da Garantia:**

Informe a data do término da garantia.

**Campo 6.8. Equipamento importado?**

Selecione a opção “Sim” caso o equipamento seja importado. Esta informação pode ser verificada na embalagem ou manual de instrução do equipamento.

Selecione a opção “Não” caso o equipamento seja nacional. Esta informação pode ser verificada na embalagem ou manual de instrução do equipamento.

**1.2.3. Bloco 7– Dados do Fabricante (seleção da opção “Não” no campo 6.8)**

Este Bloco contém os campos para preenchimento dos dados do fabricante.

Caso tenha sido realizada pesquisa nos campos anteriores, os campos deste bloco poderão ser preenchidos automaticamente.

6.8. Equipamento importado?	
<input type="radio"/> Sim <input checked="" type="radio"/> Não	
<b>7 - Dados do fabricante</b>	
7.1. Nome ou razão social do fabricante:	
<input type="text"/>	
7.2. Endereço do fabricante:	
<input type="text"/>	
7.3. Número do telefone/SAC do fabricante:	
<input type="text"/>	
7.4. UF do fabricante:	7.5. Município do fabricante:
<input type="text" value="...: Selecione ..."/>	<input type="text" value="...: Selecione ..."/>

**Campo 7.1. Nome ou razão social do fabricante:**

Informe o nome ou razão social do fabricante do equipamento.

Este campo pode ser obrigatório, dependendo da seleção do tipo de queixa técnica, no campo 3.1, da aba Motivo.

**Campo 7.2. Endereço do fabricante:**

Informe o endereço do fabricante do equipamento.

**Campo 7.3. Número do telefone/SAC do fabricante:**

Informe o número do telefone ou SAC do fabricante do equipamento.

**Campo 7.4. U.F. do fabricante:**

Selecione o estado do fabricante do equipamento. Este campo deve ser selecionado antes da seleção do campo 7.5. *Município do fabricante.*

**Campo 7.5. Município do fabricante:**

Selecione o município do fabricante do equipamento. Este campo deve ser selecionado após a seleção do campo 7.4. *UF do fabricante.*

**1.2.4. Bloco 7– Dados do Fabricante (seleção da opção “Sim” no campo 6.8)**

Este Bloco contém os campos para preenchimento dos dados do fabricante.

Caso tenha sido realizada pesquisa nos campos anteriores, os campos deste bloco poderão ser preenchidos automaticamente.

6.8. Equipamento importado? <input checked="" type="radio"/> Sim <input type="radio"/> Não	
<b>7 - Dados do importador</b>	
7.1. Nome ou razão social do importador: <input type="text"/>	
7.2. Endereço do importador: <input type="text"/>	
7.3. Número do telefone/SAC do importador: <input type="text"/>	
7.4. UF do importador: <input type="text" value="...: Selecione ..."/>	7.5. Município do importador: <input type="text" value="...: Selecione ..."/>
7.6. Nome do fabricante: <input type="text"/>	7.7. País do fabricante: <input type="text" value="...: Selecione ..."/>

**Campo 7.1. Nome ou razão social do importador:**

Informe o nome ou a razão social do importador do equipamento no Brasil. Este campo pode ser obrigatório, dependendo da seleção do tipo de queixa técnica, no campo 3.1, da aba Motivo.

**Campo 7.2. Endereço do importador:**

Informe o endereço do importador.

**Campo 7.3. Número do telefone/SAC do importador:**

Informe o número do telefone ou SAC do importador do equipamento.

**Campo 7.4. U.F. do importador:**

Selecione o estado do importador do equipamento. Este campo deve ser selecionado antes da seleção do campo 7.5. *Município do importador.*

**Campo 7.5. Município do importador:**

Selecione o município do importador do equipamento. Apenas é possível selecionar este campo após a seleção do campo 7.4. *UF do importador.*

**Campo 7.6 Nome do fabricante:**

Informe o nome ou razão social do fabricante do equipamento.

**Campo 7.7. País do Fabricante:**

Selecione o país do fabricante do produto.

**Botão Limpar dados da empresa e do produto**



Caso haja inconsistência ou incoerência no resultado da pesquisa realizada, clique nesse botão e os dados preenchidos nesta aba serão limpos para novo preenchimento.

## 1.3. Aba Paciente

Nesta aba, estão disponíveis o Bloco 8 – **Informações do paciente ou usuário** e o Bloco 9 – **Evolução do Caso**.

Motivo	Produto e Empresa	<b>Paciente</b>	Outras Informações	Pendências
--------	-------------------	-----------------	--------------------	------------

### 1.3.1. Bloco 8 – Informações do paciente ou usuário

Este Bloco contém os campos para o preenchimento dos dados do paciente ou usuário que sofreu o evento adverso

8 - Informações do paciente ou usuário		
8.1. Nome completo do paciente: (*) <input type="text"/>	8.2. Iniciais: <input type="text"/>	8.3. <input type="checkbox"/> Nome e iniciais ignorados
8.4. Nome completo da mãe do paciente: <input type="text"/>	8.5. Sexo: <input type="radio"/> Masculino <input type="radio"/> Feminino	
8.6. Raça/Cor: ...: Selecione ...	8.7. Ocupação do paciente ou usuário: ...: Selecione ...	
8.8. Data de nascimento: Dia <input type="text"/> Mês <input type="text"/> Ano <input type="text"/>	8.9. Idade: <input type="text"/> ...: Selecione ...	
8.10. Número do prontuário: <input type="text"/>	8.11. Número do Cartão SUS: <input type="text"/>	
8.12. O evento levou à internação? (*) <input type="radio"/> Sim <input type="radio"/> Não <input type="radio"/> Ignorado		
8.12.1. Hospital: <input type="text"/>		
8.12.2. UF: (*) <input type="text"/>		
8.12.3. Município: (*) <input type="text"/>		
8.12.4. Data de Internação: Dia <input type="text"/> Mês <input type="text"/> Ano <input type="text"/>		
8.12.5. Data de Alta: Dia <input type="text"/> Mês <input type="text"/> Ano <input type="text"/>		

#### Campo 8.1. Nome completo do paciente:

Informe o nome completo do paciente ou usuário. O preenchimento deste campo gera automaticamente as informações do campo 8.2. *Iniciais*.

A identificação do paciente ou usuário será mantida sob sigilo e objetiva apenas controlar possíveis duplicidades de notificações e facilitar contato com o notificador.

#### Campo 8.2. Iniciais:

Campo preenchido automaticamente pelo sistema a partir do campo 8.1, composto das iniciais do nome completo do paciente ou usuário.

Caso o campo 8.1. não tenha sido preenchido, informe as iniciais manualmente.

**Campo 8.3. Nome e Iniciais ignorados:**

Selecione quando desconhecer as informações do nome do paciente ou usuário.

O preenchimento de pelo menos um dos campos citados acima (8.1., 8.2. ou 8.3.) é obrigatório.

**Campo 8.4. Nome completo da mãe do paciente:**

Informe o nome completo da mãe do paciente ou usuário.

A identificação da mãe do paciente ou usuário será mantida sob sigilo e objetiva apenas controlar possíveis duplicidades de notificações pela identificação de homônimos.

**Campo 8.5. Sexo:**

Selecione o sexo do paciente ou usuário que sofreu o evento adverso.

**Campo 8.6. Raça/Cor:**

Selecione uma das opções para caracterizar a raça/cor do paciente ou usuário que sofreu o evento adverso:

- Branca
- Preta
- Amarela
- Parda
- Indígena
- Ignorada

**Campo 8.7. Ocupação do paciente ou usuário:**

Selecione a ocupação do paciente ou usuário que sofreu o evento adverso.

**Campo 8.8. Data de Nascimento:**

Selecione a data de nascimento do paciente ou usuário que sofreu o evento adverso.

**Campo 8.9. Idade:**

Informe a idade do paciente ou usuário no momento da ocorrência do evento. Este campo é preenchido automaticamente pelo sistema quando a data de nascimento (campo 8.12) e a data da ocorrência (campo 4.3.) forem informadas.

**Campo 8.10. Número do prontuário:**

Informe o número do prontuário do paciente, caso exista.

**Campo 8.11. Número do Cartão SUS:**

Informe o número do Cartão SUS do paciente, caso exista.

**Campo 8.12. O evento levou à internação?**

Selecione "Sim" caso o evento tenha levado à internação. Neste caso, preencher os seguintes campos:

**Campo 8.12.1. Hospital:**

Informe o nome do hospital onde o paciente foi internado.

**Campo 8.12.2. UF:**

Selecione o estado onde se localiza o hospital em que o paciente foi internado. Esse campo deve ser selecionado antes da seleção do campo 8.12.3. *Município*.

**Campo 8.12.3. Município:**

Selecione o município onde se localiza o hospital em que o paciente foi internado. Esse campo deve ser selecionado após a seleção do campo 8.12.2. *UF*.

**Campo 8.12.4. Data de Internação:**

Selecione a data de internação.

**Campo 8.12.5. Data de Alta:**

Selecione a data de alta.

Selecione “Não” caso o evento não tenha levado à internação.

Selecione “Ignorado” caso não tenha conhecimento da ocorrência de internação em decorrência do evento.

**1.3.2. Bloco 9 – Evolução do caso**

Este bloco contém as informações sobre a evolução do caso. Essas informações serão utilizadas como subsídio para classificação de gravidade.

9 - Evolução do caso
9.1. Foi a óbito? (*) <input type="radio"/> Sim <input type="radio"/> Não
9.2. Houve lesão permanente? (*) <input type="radio"/> Sim <input type="radio"/> Não
9.3. Houve lesão temporária? (*) <input type="radio"/> Sim <input type="radio"/> Não

**Campo 9.1. Foi a óbito?**

Campo de preenchimento obrigatório.

Selecione “Sim” se o paciente evoluiu para óbito. Neste caso, a mensagem “Copiar tal qual consta na Declaração de Óbito” e os seguintes campos são apresentados:

**Campo 9.1.1. Data:**

Selecione a data em que ocorreu o óbito.

**Campo 9.1.2. Causa da morte - Campo 49 da declaração de óbito:**

Informe a causa do óbito. Este campo possibilita a inclusão de até quatro causas do óbito (itens a, b, c, d).

Selecione “Não” caso não tenha ocorrido o óbito do paciente.



### **Campo 9.2. Houve lesão permanente?**

Campo de preenchimento obrigatório.

Selecione “Sim” caso tenha ocorrido lesão permanente decorrente do evento. Neste caso, o seguinte campo é apresentado:

#### **Campo 9.2.1 Descrever:**

Descreva a lesão permanente.

Selecione “Não” caso não tenha ocorrido lesão permanente decorrente do evento.

### **Campo 9.3. Houve lesão temporária?**

Campo de preenchimento obrigatório.

Selecione “Sim” caso tenha ocorrido lesão temporária decorrente do evento. Neste caso, o seguinte campo é apresentado:

#### **Campo 9.3.1 Descrever:**

Campo alfa-numérico. Informe a descrição da lesão temporária.

Selecione “Não” caso não tenha ocorrido lesão temporária decorrente do evento.

## **1.4. Aba Outras Informações**

Nesta aba consta apenas o Bloco 10 – **Outras informações importantes**



### **1.4.1. Bloco 10 – Outras informações importantes**

Este Bloco contém campos para o preenchimento de informações adicionais referentes ao equipamento que motivou a notificação.

10 - Outras informações importantes	
10.1. A utilização do produto seguiu as instruções do fabricante?	<input type="radio"/> Sim <input type="radio"/> Não
10.2. O produto apresenta alterações (suspeita de desvio de qualidade)?	<input type="radio"/> Sim <input type="radio"/> Não <input type="radio"/> Ignorado
10.3. Local de aquisição do produto:	<input type="text" value="...: Selecione ..."/>
10.4. Possui nota fiscal da compra do produto?	<input type="radio"/> Sim <input type="radio"/> Não <input type="radio"/> Ignorado
10.5. Houve comunicação à indústria/distribuidor?	<input type="radio"/> Sim <input type="radio"/> Não <input type="radio"/> Ignorado
10.6. Foram adotadas outras providências após a identificação do problema?	<input type="radio"/> Sim <input type="radio"/> Não
10.7. Existem amostras íntegras para a coleta?	<input type="radio"/> Sim <input type="radio"/> Não
10.8. Existem rótulos do produto para a coleta?	<input type="radio"/> Sim <input type="radio"/> Não
10.9. Observações:	<input type="text"/>

#### **Campo 10.1. A utilização do produto seguiu as instruções do fabricante?**

Selecione a opção “Sim” caso o equipamento tenha sido utilizado conforme as orientações constantes no manual de instrução do equipamento.

Selecione a opção “Não” caso o equipamento não tenha sido utilizado conforme as orientações constantes no manual de instrução do equipamento.

#### **Campo 10.2. O equipamento apresenta alterações (suspeita de desvio de qualidade)?**

Selecione a opção “Sim” caso observe alterações no produto ou suspeite de algum desvio da qualidade. Neste caso, selecione uma ou mais opções apresentadas.

10.2. O equipamento apresenta alterações (suspeita de desvio de qualidade)?

- Sim  Não  Ignorado
- Não liga
  - Não desliga
  - Fornece dados incorretos
  - Alarme dispara frequentemente
  - Alarme não dispara
  - Travamento
  - Choque elétrico
  - Falhas de bateria
  - Vazamento
  - Superaquecimento
  - Fuga de corrente
  - Quebra frequente de partes e/ou peças do equipamento
  - Não desempenha a função requerida
  - Obstrução da saída de gás
  - Outra

### Campo 10.3. Local de aquisição do equipamento:

Selecione o local de aquisição do produto de acordo com as opções disponíveis.

10.3. Local de aquisição do equipamento:

...: Selecione ...

- ...: Selecione ...
- Distribuidora
- Drogeria
- Farmácias
- Importadora
- Indústria
- Internet
- Mercado
- Representante
- Televenda
- Estabelecimento de saúde pública

Caso o local de aquisição não esteja disponível na lista, selecione a opção "Outro". Nesta situação, o campo 10.3.1 *Citar* é apresentado para descrição.

10.3. Local de aquisição do equipamento:

Outro

10.3.1. Citar:

### Campo 10.4. Forma de aquisição do equipamento?

Selecione a forma de aquisição do equipamento. Para este campo estão disponíveis as seguintes opções:

10.4. Forma de aquisição do equipamento:

...: Selecione ...

- ...: Selecione ...
- Ignorado
- Compra direta
- Doação
- Licitação
- Outro

Caso a forma de aquisição não esteja disponível na lista, selecione a opção "Outro". Nesta situação, o campo 10.4.1 *Citar* é apresentado para descrição.

10.4. Forma de aquisição do equipamento:

Outro

10.4.1. Citar:

### Campo 10.5. Quem faz a manutenção do equipamento?

Selecione o responsável pela realização de manutenção do equipamento. Para este campo estão disponíveis as seguintes opções:

10.5. Quem faz a manutenção do equipamento?

...: Selecione ...

- ...: Selecione ...
- Equipe própria
- Serviço Terceirizado
- Ambos
- Outro

Caso não esteja disponível na lista, selecione a opção "Outro". Nesta situação, o campo 10.5.1 *Citar* é apresentado para descrição.

10.5. Quem faz a manutenção do equipamento?

Outro

10.5.1. Citar:

### Campo 10.6. Histórico recente das três últimas manutenções

Informe as datas das três últimas manutenções e descreva o procedimento realizado. Por exemplo: se manutenção preventiva ou corretiva e os procedimentos realizados, como troca de peças, por exemplo.

### Campo 10.7. Possui nota fiscal de compra do equipamento?

Selecione a opção “Sim” caso possua a nota fiscal de compra do equipamento disponível para apresentação ao órgão do SNVS responsável, se necessário. Neste caso, será apresentada a mensagem “Mantenha-a em seu poder” em vermelho.

10.4. Possui nota fiscal da compra do produto?

Sim  Não  Ignorado

Mantenha-a em seu poder.

Selecione a opção “Não” caso não possua a nota fiscal de compra do equipamento.

Selecione a opção “Ignorado” caso não tenha conhecimento da disponibilidade da nota fiscal de compra do equipamento.

### Campo 10.8. Houve comunicação à indústria/distribuidor?

Selecione a opção “Sim” caso a queixa técnica também tenha sido informada à indústria/distribuidor. Neste caso, são apresentadas as opções mostradas abaixo para seleção da forma de comunicação. Após seleção é apresentada a mensagem “Caso possua comprovante, mantenha-o em seu poder e somente entregue à Vigilância Sanitária quando solicitado”.

8.4. Houve comunicação à indústria/distribuidor?

Sim  Não  Ignorado

- Por correio
- Por e-Mail
- Por Fax
- Por representante
- Por Telefone
- Outra Forma

Caso possua comprovante, mantenha-o em seu poder e somente entregue à Vigilância Sanitária quando solicitado.

Caso a forma de comunicação não esteja disponível na lista, selecione a opção “Outra forma”. Nesta situação, o campo 8.4.1 Citar é apresentado para descrição.

8.4.1. Citar:

Selecione a opção “Não” caso a queixa técnica não tenha sido comunicada à indústria/distribuidor.

Selecione a opção “Ignorado” caso não tenha conhecimento da comunicação da queixa técnica à indústria/distribuidor.

### Campo 10.9. Foram adotadas outras providências após a identificação do problema?

Selecione a opção “Sim” caso tenham sido adotadas outras providências, além das descritas neste formulário de notificação. Nesta situação, o campo 10.9.1 Citar é apresentado para descrição.

10.9.1. Citar:

Selecione a opção “Não” caso não tenham sido adotadas outras providências, além das descritas neste formulário de notificação.

### Campo 10.10. Observações

Caso considere relevante alguma informação ainda não relatada, informe neste campo.

## 1.5. Aba Pendências

A qualquer momento as pendências poderão ser verificadas selecionando esta aba.

Nesta Aba, são apontadas as pendências quanto à violação de regras e obrigadoriedades no preenchimento do formulário de notificação. Está disponível o Bloco – **Pendências**.

Motivo	Produto e Empresa	Paciente	Outras Informações	<b>Pendências</b>
--------	-------------------	----------	--------------------	-------------------

### 1.5.1. Bloco – Pendências

Este Bloco apresenta os campos que necessitam de verificação para o notificador providenciar a correção ou preenchimento antes do envio.

Pendências
Os seguintes campos necessitam verificação:
<b>Motivo:</b>
Data da identificação: Campo obrigatório não preenchido.
Descreva detalhadamente a Queixa Técnica/Evento Adverso: Campo obrigatório não preenchido.
Município: Campo obrigatório não preenchido.
Selecione uma das opções: Campo obrigatório não preenchido.

Após o preenchimento do formulário, caso não haja pendências, será apresentada a mensagem “A Notificação foi verificada e não existe pendência”.

Para enviar a notificação, clique no botão  . O certificado de envio será apresentado:

**Notificação Enviada**

**Número da Notificação:** 2006.10.001908  
**Data da notificação:** 25/10/2006  
**Notificador:** bruno TecnicoSNVS  
**Produto motivo da notificação:** Equipamento

Caro(a) Notificador(a),  
sua notificação foi recebida com sucesso e será analisada pelos técnicos da Vigilância Sanitária.  
Atenciosamente,

NOTIVISA - Sistema Nacional de Notificações para a Vigilância Sanitária

Para imprimir o formulário preenchido, clique em  e para imprimir o certificado, clique em  .

## Anexo IV

# Formulário para Notificação de Queixa Técnica de Equipamento Médico-Hospitalar

# Manual do Usuário



Notivisa – Sistema Nacional de Notificações  
para a Vigilância Sanitária

## Anexo 9

**Formulário para Notificação de Queixa Técnica de  
Equipamento Médico-Hospitalar**



## Sumário

1. FORMULÁRIO PARA NOTIFICAÇÃO DE QUEIXA TÉCNICA DE EQUIPAMENTO MÉDICO-HOSPITALAR .....	3
<b>1.1. ABA MOTIVO .....</b>	<b>3</b>
1.1.1. Bloco 3 – Tipo de Queixa Técnica .....	3
1.1.2. Bloco 4 – Queixa Técnica .....	4
<b>1.2. ABA PRODUTO E EMPRESA .....</b>	<b>6</b>
1.2.1. Bloco 5 – Produto e empresa .....	6
1.2.2. Bloco 6– Dados do Produto .....	7
1.2.3. Bloco 7– Dados do Fabricante (seleção da opção “Não” no campo 6.8).8	
1.2.4. Bloco 7– Dados do Fabricante (seleção da opção “Sim” no campo 6.8).9	
<b>1.3. ABA OUTRAS INFORMAÇÕES .....</b>	<b>10</b>
1.3.1. Bloco 10 – Outras informações importantes .....	10
<b>1.4. ABA PENDÊNCIAS .....</b>	<b>14</b>
1.4.1 Bloco – Pendências .....	14

## 1. Formulário para Notificação de Queixa Técnica de Equipamento Médico-Hospitalar

Para Queixa Técnica de Equipamento Médico-Hospitalar (marcapasso, ventilador pulmonar, bomba de infusão, etc.) estão disponíveis as seguintes abas:

- Motivo;
- Produto e Empresa;
- Outras informações;
- Pendências.

### 1.1. Aba Motivo

Motivo	Produto e Empresa	Outras Informações	Pendências
--------	-------------------	--------------------	------------

#### 1.1.1. Bloco 3 – Tipo de Queixa Técnica

Este Bloco contém as opções para caracterização da queixa técnica a ser notificada.

3 - Tipo de Queixa Técnica
3.1. Selecione uma das opções: (*)
...: Selecione ...

Selecione uma das opções disponíveis no campo 3.1. Este campo é de preenchimento obrigatório.

Opção	Descrição
Produto com suspeita de desvio da qualidade	Selecione quando houver suspeita de afastamento dos parâmetros de qualidade estabelecidos para um produto ou processo.
Produto com suspeita de estar sem Registro	Selecione quando houver suspeita de que o produto não está registrado ou está com o registro vencido. O número de registro na Anvisa/Ministério da Saúde (MS) deve estar impresso no rótulo ou embalagem do produto.
Suspeita de empresa sem autorização de funcionamento (AFE)	Selecione quando houver suspeita de que a empresa fabricante ou importadora do produto não tem autorização de funcionamento (AFE), que é emitida pela esfera competente do Sistema Nacional de Vigilância Sanitária (SNVS).
Suspeita de produto falsificado	Selecione quando houver suspeita de que o produto foi falsificado, ou seja, simula a identidade de outro produto já legalmente registrado.
Suspeita de outras práticas irregulares	Selecione para indicar suspeita de outras práticas não citadas nas opções acima.

Quando a opção “Produto com suspeita de desvio de qualidade” for selecionada, assinale uma das opções do campo “3.1.1. O produto apresenta alterações?”, conforme segue:

3 - Tipo de Queixa Técnica	
3.1. Selecione uma das opções: (*)	<input type="text" value="Produto com suspeita de desvio da qualidade"/>
3.1.1. O produto apresenta alterações?	
<input type="checkbox"/> Não liga	
<input type="checkbox"/> Não desliga	
<input type="checkbox"/> Fornece dados incorretos	
<input type="checkbox"/> Alarme dispara frequentemente	
<input type="checkbox"/> Alarme não dispara	
<input type="checkbox"/> Travamento	
<input type="checkbox"/> Choque elétrico	
<input type="checkbox"/> Falhas de bateria	
<input type="checkbox"/> Vazamento	
<input type="checkbox"/> Superaquecimento	
<input type="checkbox"/> Fuga de corrente	
<input type="checkbox"/> Quebra frequente de partes e/ou peças do equipamento	
<input type="checkbox"/> Não desempenha a função requerida	
<input type="checkbox"/> Obstrução da saída de gás	
<input type="checkbox"/> Outra	

Caso a alteração não esteja disponível na lista, selecione a opção “Outra”. Nesta situação, o campo 3.1.2. *Citar*: é apresentado para descrição.

3.1.2. *Citar*:

Quando a opção “Suspeita de outras práticas irregulares” for assinalada, o campo 3.1.2. *Citar*: é apresentado para descrição.

3.1. Selecione uma das opções: (\*)

3.1.2. *Citar*:

### 1.1.2. Bloco 4 – Queixa Técnica

Este Bloco contém os campos para preenchimento de dados da identificação e descrição da Queixa Técnica.

4 - Queixa Técnica		
4.1. Descreva detalhadamente a Queixa Técnica : (*)		
<div style="border: 1px solid #ccc; height: 40px;"></div>		
4.3. Data da identificação: (*)		
Dia <input type="text"/> Mês <input type="text"/> Ano <input type="text"/>		
4.4. Local da identificação:	4.5. Endereço do local da identificação:	
<input type="text" value="Estabelecimento de saúde"/>	<input type="text" value="515N - Edifício ANVISA"/>	
4.4.1. Nome do estabelecimento de saúde:		
<input type="text" value="Empresa de Teste Ltda"/>		
4.4.2. CNPJ do estabelecimento de saúde:		
<input type="text" value="33.683.202/0001-34"/>		
4.6. País: (*)	4.7. UF: (*)	4.8. Município: (*)
<input type="text" value="BRASIL"/>	<input type="text" value="...: Selecione ..."/>	<input type="text" value="...: Selecione ..."/>

#### **Campo 4.1. Descreva detalhadamente a Queixa Técnica:**

Campo de preenchimento obrigatório. Descreva detalhadamente a queixa técnica incluindo os motivos da notificação.

#### **Campo 4.3. Data da identificação:**

Campo de preenchimento obrigatório. Selecione a data da identificação do problema (queixa técnica).

#### **Campo 4.4. Local da identificação:**

Selecione o local onde o problema foi identificado dentre as seguintes opções:

- Estabelecimento de Saúde;
- Residência;
- Outros.

Caso selecione a opção “Estabelecimento de Saúde”, preencha os seguintes campos:

- 4.4.1. Nome do estabelecimento de saúde:
- 4.4.2. CNPJ do estabelecimento de saúde:

Caso selecione a opção “Outro”, preencha o seguinte campo

- 4.4.3. Citar:

#### **Campo 4.5. Endereço do local da identificação:**

Informe o endereço do local da identificação do problema (queixa técnica).

#### **Campo 4.6. País:**

Campo de preenchimento obrigatório. Selecione o país onde foi identificado o problema (queixa técnica). A opção “Brasil” está pré-selecionada.

#### Campo 4.7. UF:

Selecione o estado onde foi identificado o problema. Este campo deve ser selecionado antes da seleção do campo 4.8. *Município*. Se o país selecionado no campo 4.6 for “Brasil”, este campo é de preenchimento obrigatório.

#### Campo 4.8. Município:

Selecione o município onde foi identificado o problema. Só é possível selecionar este campo após a seleção do campo 4.7. *UF*. Se o país selecionado no campo 4.6 for “Brasil”, este campo é de preenchimento obrigatório.

## 1.2. Aba Produto e Empresa

Nesta Aba, estão disponíveis os seguintes Blocos: **Produto e Empresa**, **Dados do Produto** e **Dados do Fabricante ou Importador**.



### 1.2.1. Bloco 5 – Produto e empresa

Este Bloco contém os campos para preenchimento do Número do Registro do produto na ANVISA/MS e do CNPJ da empresa fabricante ou importadora. Para facilitar o preenchimento de dados desta aba, opcionalmente, o notificador poderá realizar pesquisa no banco de dados da Anvisa.

5 - Produto e Empresa	
5.1. Número do registro ou cadastro na ANVISA/MS: <input type="text"/>	5.2. CNPJ da empresa fabricante ou importador: <input type="text"/>
<input type="button" value="Pesquisar"/>	<input type="button" value="Pesquisar"/>

Dependendo da pesquisa realizada, os campos do Bloco 5 – **Produto e Empresa**, Bloco 6 - **Dados do produto** e Bloco 7 – **Dados do Fabricante ou Importador** são preenchidos automaticamente.

Estes Campos podem ser preenchidos automaticamente por pesquisa realizada em campos do Bloco 6 – **Dados do Produto**.

#### Campo 5.1. Número do registro ou cadastro na ANVISA/MS:

Campo numérico com espaço para até 13 dígitos. Informe o número do registro do equipamento na ANVISA/MS, impresso no rótulo/embalagem ou no próprio equipamento.

Após preenchimento, clique no botão “pesquisar”.

Caso os dados do equipamento sejam encontrados no banco de dados da Anvisa, alguns campos referentes ao equipamento e empresa são preenchidos automaticamente. Estes dados devem ser verificados e, caso não confirmem com o produto, devem ser alterados manualmente.

#### Campo 5.2. CNPJ da empresa fabricante ou importador:

Campo numérico com espaço para até 14 dígitos. Informe o CNPJ da empresa fabricante ou importadora impresso no equipamento.

Após preenchimento, clique no botão “pesquisar”.

Caso os dados da empresa sejam encontrados no banco de dados da Anvisa, alguns campos referentes ao produto e empresa são preenchidos automaticamente. Estes dados devem ser verificados e, caso não confirmem com o produto, devem ser alterados manualmente.

### 1.2.2. Bloco 6– Dados do Produto

Este Bloco contém os campos para o preenchimento dos dados referentes ao produto, que poderão ser preenchidos automaticamente como resultado da pesquisa realizada no Bloco 5.

5 - Produto e Empresa	
5.1. Número do registro ou cadastro na ANVISA/MS: <input type="text"/>	5.2. CNPJ da empresa fabricante ou importador: <input type="text"/>
<input type="button" value="Pesquisar"/>	<input type="button" value="Pesquisar"/>
6 - Dados do Produto	
6.1. Nome comercial do equipamento: <input type="text"/>	
<input type="button" value="Pesquisar"/>	
6.2. Equipamento: ...Selecione:.. <input type="button" value="v"/>	
6.3. Modelo do equipamento: <input type="text"/>	6.4. Número de série de fabricação: <input type="text"/>
6.5. Data de aquisição: Dia <input type="button" value="v"/> Mês <input type="button" value="v"/> Ano <input type="button" value="v"/>	6.6. Data de instalação: Dia <input type="button" value="v"/> Mês <input type="button" value="v"/> Ano <input type="button" value="v"/>
6.7. Data do término da garantia: Dia <input type="button" value="v"/> Mês <input type="button" value="v"/> Ano <input type="button" value="v"/>	
6.8. Equipamento importado? <input type="radio"/> Sim <input type="radio"/> Não	

#### Campo 6.1. Nome comercial do equipamento:

Campo para identificação do equipamento. Preencha com o nome comercial do equipamento constante na sua embalagem ou manual de instrução.

Para a realização da pesquisa a partir deste campo, pelo menos 3 (três) caracteres devem ser informados.

Após clicar em “Pesquisar”, selecione uma das opções da lista apresentada. Caso os dados confirmem com o produto em mãos, clique no botão  e preencha todos os campos manualmente.

Este campo pode ser obrigatório, dependendo da seleção do tipo de queixa técnica, no campo 3.1, da aba Motivo.

Caso tenha sido realizada pesquisa pelo CNPJ, no campo 5.2, este campo poderá apresentar lista com equipamentos fabricados pela empresa em questão para seleção do equipamento que motivou a notificação.

#### Campo 6.2. Equipamento:

Campo para seleção da categoria do equipamento. Este campo não pode ser preenchido manualmente.

Após a realização de pesquisa nos campos 5.1., 5.2. ou 6.1., o sistema poderá carregar automaticamente uma tabela para a seleção da categoria do equipamento. Caso não haja opções para seleção, desconsidere este campo.

**Campo 6.3. Modelo do Equipamento:**

Informe o modelo do equipamento.

**Campo 6.4. Número de série de fabricação:**

Informe o número de série de fabricação.

**Campo 6.5. Data de aquisição:**

Informe a data de aquisição do equipamento.

**Campo 6.6. Data de instalação:**

Informe a data de instalação de instalação do equipamento no serviço de saúde.

**Campo 6.7. Data do Término da Garantia:**

Informe a data do término da garantia.

**Campo 6.8. Equipamento importado?**

Selecione a opção “Sim” caso o equipamento seja importado. Esta informação pode ser verificada na embalagem ou manual de instrução do equipamento.

Selecione a opção “Não” caso o equipamento seja nacional Esta informação pode ser verificada na embalagem ou manual de instrução do equipamento.

**1.2.3. Bloco 7– Dados do Fabricante (seleção da opção “Não” no campo 6.8)**

Este Bloco contém os campos para preenchimento dos dados do fabricante.

Caso tenha sido realizada pesquisa nos campos anteriores, os campos deste bloco poderão ser preenchidos automaticamente.

6.8. Equipamento importado? <input type="radio"/> Sim <input checked="" type="radio"/> Não	
<b>7 - Dados do fabricante</b>	
7.1. Nome ou razão social do fabricante: <input type="text"/>	
7.2. Endereço do fabricante: <input type="text"/>	
7.3. Número do telefone/SAC do fabricante: <input type="text"/>	
7.4. UF do fabricante: <input type="text" value="...: Seleccione ..."/>	7.5. Município do fabricante: <input type="text" value="...: Seleccione ..."/>

**Campo 7.1. Nome ou razão social do fabricante:**

Informe o nome ou razão social do fabricante do equipamento.

Este campo pode ser obrigatório, dependendo da seleção do tipo de queixa técnica, no campo 3.1, da aba Motivo.

**Campo 7.2. Endereço do fabricante:**

Informe o endereço do fabricante do equipamento.

**Campo 7.3. Número do telefone/SAC do fabricante:**

Informe o número do telefone ou SAC do fabricante do equipamento.

**Campo 7.4. U.F. do fabricante:**

Selecione o estado do fabricante do equipamento. Este campo deve ser selecionado antes da seleção do campo 7.5. *Município do fabricante.*

**Campo 7.5. Município do fabricante:**

Selecione o município do fabricante do equipamento. Este campo deve ser selecionado após a seleção do campo 7.4. *UF do fabricante.*

**1.2.4. Bloco 7– Dados do Fabricante (seleção da opção “Sim” no campo 6.8)**

Este Bloco contém os campos para preenchimento dos dados do fabricante.

Caso tenha sido realizada pesquisa nos campos anteriores, os campos deste bloco poderão ser preenchidos automaticamente.

6.8. Equipamento importado? <input checked="" type="radio"/> Sim <input type="radio"/> Não	
<b>7 - Dados do importador</b>	
7.1. Nome ou razão social do importador: <input type="text"/>	
7.2. Endereço do importador: <input type="text"/>	
7.3. Número do telefone/SAC do importador: <input type="text"/>	
7.4. UF do importador: ...: Seleccione ...	7.5. Município do importador: ...: Seleccione ...
7.6. Nome do fabricante: <input type="text"/>	7.7. País do fabricante: ...: Seleccione ...

**Campo 7.1. Nome ou razão social do importador:**



Informe o nome ou a razão social do importador do equipamento no Brasil. Este campo pode ser obrigatório, dependendo da seleção do tipo de queixa técnica, no campo 3.1, da aba Motivo.

**Campo 7.2. Endereço do importador:**

Informe o endereço do importador.

**Campo 7.3. Número do telefone/SAC do importador:**

Informe o número do telefone ou SAC do importador do equipamento.

**Campo 7.4. U.F. do importador:**

Selecione o estado do importador do equipamento. Este campo deve ser selecionado antes da seleção do campo 7.5. *Município do importador.*

**Campo 7.5. Município do importador:**

Selecione o município do importador do equipamento. Apenas é possível selecionar este campo após a seleção do campo 7.4. *UF do importador.*

**Campo 7.6 Nome do fabricante:**

Informe o nome ou razão social do fabricante do equipamento.

**Campo 7.7. País do Fabricante:**

Selecione o país do fabricante do produto.

**Botão Limpar dados da empresa e do produto**

Caso haja inconsistência ou incoerência no resultado da pesquisa realizada, clique nesse botão e os dados preenchidos nesta aba serão limpos para novo preenchimento.

## 1.3. ABA OUTRAS INFORMAÇÕES

Nesta Aba, está disponível o **Bloco 10 – Outras informações importantes**



### 1.3.1. Bloco 10 – Outras informações importantes

Este Bloco contém os campos para o preenchimento de informações adicionais referentes ao equipamento que motivou a notificação.

10 - Outras informações importantes	
10.1. A utilização do produto seguiu as instruções do fabricante?	<input type="radio"/> Sim <input type="radio"/> Não
10.3. Local de aquisição do equipamento:	<input type="text" value="...: Selecione :..."/>
10.4. Forma de aquisição do equipamento:	<input type="text" value="...: Selecione :..."/>
10.5. Quem faz a manutenção do equipamento?	<input type="text" value="...: Selecione :..."/>
10.6. Histórico recente das três últimas manutenções: <i>Informar as respectivas datas (dd/mm/aaaa) e descrever os procedimentos adotados em cada uma.</i>	<input type="text"/>
10.7. Possui nota fiscal da compra do equipamento?	<input type="radio"/> Sim <input type="radio"/> Não <input type="radio"/> Ignorado
10.8. Houve comunicação à indústria/distribuidor?	<input type="radio"/> Sim <input type="radio"/> Não <input type="radio"/> Ignorado
10.9. Foram adotadas outras providências após a identificação do problema?	<input type="radio"/> Sim <input type="radio"/> Não
10.10. Observações:	<input type="text"/>

**Campo 10.1. A utilização do produto seguiu as instruções do fabricante?**

Selecione a opção “Sim” caso o equipamento tenha sido utilizado conforme as orientações constantes no manual de instrução do equipamento.

Selecione a opção “Não” caso o equipamento não tenha sido utilizado conforme as orientações constantes no manual de instrução do equipamento.

**Campo 10.3. Local de aquisição do equipamento:**

Selecione o local de aquisição do produto de acordo com as opções disponíveis.

10.3. Local de aquisição do equipamento:

... Seleccione ...

- ... Seleccione ...
- Distribuidora
- Drogeria
- Farmácias
- Importadora
- Indústria
- Internet
- Mercado
- Representante
- Televenda
- Estabelecimento de saúde pública

Caso o local de aquisição não esteja disponível na lista, selecione a opção “Outro”. Nesta situação, o campo *10.3.1 Citar* é apresentado para descrição.

10.3. Local de aquisição do equipamento:

Outro

10.3.1. Citar:

#### Campo 10.4. Forma de aquisição do equipamento?

Selecione a forma de aquisição do equipamento. Para este campo estão disponíveis as seguintes opções:

10.4. Forma de aquisição do equipamento:

... Seleccione ...

- ... Seleccione ...
- Ignorado
- Compra direta
- Doação
- Licitação
- Outro

Caso a forma de aquisição não esteja disponível na lista, selecione a opção “Outro”. Nesta situação, o campo *10.4.1 Citar* é apresentado para descrição.

10.4. Forma de aquisição do equipamento:

Outro

10.4.1. Citar:

#### Campo 10.5. Quem faz a manutenção do equipamento?

Selecione o responsável pela realização de manutenção do equipamento. Para este campo estão disponíveis as seguintes opções:

10.5. Quem faz a manutenção do equipamento?

... Seleccione ...

- ... Seleccione ...
- Equipe própria
- Serviço Terceirizado
- Ambos
- Outro

Caso não esteja disponível na lista, selecione a opção “Outro”. Nesta situação, o campo *10.5.1 Citar* é apresentado para descrição.

10.5. Quem faz a manutenção do equipamento?

Outro

10.5.1. Citar:

### Campo 10.6. Histórico recente das três últimas manutenções

Informe as datas das três últimas manutenções e descreva o procedimento realizado. Por exemplo: se manutenção preventiva ou corretiva e os procedimentos realizados, como troca de peças, por exemplo.

### Campo 10.7. Possui nota fiscal de compra do equipamento?

Selecione a opção “Sim” caso possua a nota fiscal de compra do equipamento disponível para apresentação ao órgão do SNVS responsável, se necessário. Neste caso, será apresentada a mensagem “Mantenha-a em seu poder” em vermelho.

10.4. Possui nota fiscal da compra do produto?

Sim  Não  Ignorado

Mantenha-a em seu poder.

Selecione a opção “Não” caso não possua a nota fiscal de compra do equipamento.

Selecione a opção “Ignorado” caso não tenha conhecimento da disponibilidade da nota fiscal de compra do equipamento.

### Campo 10.8. Houve comunicação à indústria/distribuidor?

Selecione a opção “Sim” caso a queixa técnica também tenha sido informada à indústria/distribuidor. Neste caso, são apresentadas as opções mostradas abaixo para seleção da forma de comunicação. Após seleção é apresentada a mensagem “Caso possua comprovante, mantenha-o em seu poder e somente entregue à Vigilância Sanitária quando solicitado”.

8.4. Houve comunicação à indústria/distribuidor?

Sim  Não  Ignorado

Por correio

Por e-Mail

Por Fax

Por representante

Por Telefone

Outra Forma

Caso possua comprovante, mantenha-o em seu poder e somente entregue à Vigilância Sanitária quando solicitado.

Caso a forma de comunicação não esteja disponível na lista, selecione a opção “Outra forma”. Nesta situação, o campo 8.4.1 Citar é apresentado para descrição.

8.4.1. Citar:

Selecione a opção “Não” caso a queixa técnica não tenha sido comunicada à indústria/distribuidor.

Selecione a opção “Ignorado” caso não tenha conhecimento da comunicação da queixa técnica à indústria/distribuidor.

### Campo 10.9. Foram adotadas outras providências após a identificação do problema?

Selecione a opção “Sim” caso tenham sido adotadas outras providências, além das descritas neste formulário de notificação. Nesta situação, o campo 10.9.1 Citar é apresentado para descrição.

10.9.1. Citar:

Selecione a opção “Não” caso não tenham sido adotadas outras providências, além das descritas neste formulário de notificação.

### Campo 10.10. Observações

Caso considere relevante alguma informação ainda não relatada, informe neste campo.

## 1.4. Aba Pendências

A qualquer momento as pendências poderão ser verificadas selecionando esta aba.

Nesta Aba, são apontadas as pendências quanto à violação de regras e obrigtoriedades no preenchimento do formulário de notificação. Está disponível o Bloco – **Pendências**.

Motivo	Produto e Empresa	Outras Informações	<b>Pendências</b>
--------	-------------------	--------------------	-------------------

### 1.4.1 Bloco – Pendências

Este Bloco apresenta os campos que necessitam de verificação para o notificador providenciar a correção ou preenchimento antes do envio.

Pendências
Os seguintes campos necessitam verificação:
<b>Motivo:</b>
Data da identificação: Campo obrigatório não preenchido.
Descreva detalhadamente a Queixa Técnica/Evento Adverso: Campo obrigatório não preenchido.
Município: Campo obrigatório não preenchido.
Selecione uma das opções: Campo obrigatório não preenchido.

Após o preenchimento do formulário, caso não haja pendências, será apresentada a mensagem “A Notificação foi verificada e não existe pendência”.

Para enviar a notificação, clique no botão  . O certificado de envio será apresentado:

**Notificação Enviada**

**Número da Notificação:** 2006.10.001908  
**Data da notificação:** 25/10/2006  
**Notificador:** bruno TecnicoSNVS  
**Produto motivo da notificação:** Equipamento

Caro(a) Notificador(a),  
sua notificação foi recebida com sucesso e será analisada pelos técnicos da Vigilância Sanitária.  
Atenciosamente,

NOTIVISA - Sistema Nacional de Notificações para a Vigilância Sanitária

Para imprimir o formulário preenchido, clique em  e para imprimir o certificado, clique em  .