

## **AVALIAÇÃO DO SENTIMENTO DE PRODUTORES E PROFISSIONAIS DE SAÚDE ANIMAL DO ESTADO DE MATO GROSSO EM RELAÇÃO AO SISTEMA DE VIGILÂNCIA EPIDEMIOLÓGICA PARA FEBRE AFTOSA**

**Helena Martos Romboli**

**Fernando Ferreira**

UNIVERSIDADE DE SÃO PAULO  
FACULDADE DE MEDICINA VETERINÁRIA E ZOOTECNIA (FMVZ/USP)

[helenamartosromboli@usp.br](mailto:helenamartosromboli@usp.br)

### **Objetivos**

Utilizar técnicas de inteligência artificial para avaliar o sentimento de produtores, servidores do INDEA, revendedores de produtores, servidores do INDEA, revendedores de produtos agropecuários e veterinários do serviço privado em relação a diferentes aspectos do sistema de vigilância epidemiológica da febre aftosa no estado do Mato Grosso.

Verificar a aplicabilidade dessas ferramentas para promover o engajamento desses atores ao sistema de vigilância para febre aftosa.

### **Métodos e Procedimentos**

O Laboratório de Epidemiologia e Bioestatística (LEB), do departamento de Medicina Veterinária Preventiva e Saúde Única Animal (VPS) da FMVZ/USP, vem trabalhando com a avaliação do sistema de vigilância para FA no estado do Mato Grosso. As pesquisas foram realizadas entre 11/09/2020 a 15/02/2021 em diversos municípios do estado de Mato Grosso. Para realizar as pesquisas, o projeto foi apresentado a Equipe Gestora Estadual do Plano Estratégico 2017- 2026 do Programa Nacional de Erradicação e Prevenção da Febre Aftosa, no âmbito do Instituto de Defesa Agropecuária do Estado do Mato Grosso, a qual é constituída por organizações públicas e privadas (MATO GROSSO, 2018), cujo apoio possibilitou a contratação pelo Fundo

Emergencial de Saúde Animal, da empresa terceirizada com competência e experiência requeridas para o projeto, Vetor Pesquisa, com o intuito de evitar o viés do entrevistador pertencer ao SVO. A empresa participou da fase de planejamento, elaboração dos questionários e roteiros, execução da pesquisa a campo e elaboração de relatórios com a análise exploratória dos dados e percepções das entrevistas qualitativas (VETOR, 2021). Ao final da execução dos trabalhos, forneceu os bancos de dados, os quais são utilizados para as análises.

Os roteiros para as entrevistas em profundidade foram idealizados para captar percepções que não seriam possíveis somente com a aplicação do questionário estruturado, tais quais os fatores que podem influenciar no processo de notificação das doenças e as crenças e os comportamentos relacionados com o sistema de vigilância.

Para a construção dos roteiros os seguintes tópicos foram abordados: informações sociodemográficas, avaliação dos sistemas de informações disponibilizados pelo INDEA, adoção das medidas de biosseguridade nos estabelecimentos rurais, conhecimentos sobre a febre aftosa, percepções sobre a notificação de doenças e das etapas envolvidas, suspensão da vacinação contra a febre aftosa e imagem institucional do INDEA e Fundo Emergencial de Saúde Animal (FESA – MT). Foram estabelecidos critérios para seleção do público para participar das entrevistas:

produtores rurais ou responsáveis pelo manejo nos estabelecimentos rurais com bovinos; médicos veterinários autônomos que trabalham com animais de produção, e revendas com comércio de fármacos e biológicos destinados à produção animal.

Os roteiros para as entrevistas, após finalizados, foram testados a campo através da aplicação de entrevistas piloto, onde foi observado se as perguntas eram de fácil entendimento, principalmente para o público de baixa escolaridade, se as respostas não estavam sendo coletadas de forma enviesada e observado a ordem das questões para não induzir respostas.

As equipes responsáveis pelas entrevistas, foram capacitadas pela empresa, através de treinamento prévio.

A Tabela 2 apresenta a descrição amostral de cada população e o tamanho da amostra obtida. Como se trata de aplicação de pesquisa qualitativa cujo processo é mais complexo e demorado, o tamanho da amostra foi estimado considerando a capacidade de coleta no período de realização do projeto.

Tabela 2: Descrição das amostras de pesquisa, considerando o público-alvo, universo, amostra e qualidade. Mato Grosso, 2020.

Público Alvo	Universo	Amostra qualitativa
Servidores do INDEA	821	20
Produtores rurais	108.079	40
Médico veterinário	1.255	20
Revendas de produtos veterinários	1.006	20

Fonte: SCHMIDT, Ana Carolina. AVALIAÇÃO E PROPOSTA DE REESTRUTURAÇÃO DO SISTEMA DE VIGILÂNCIA EPIDEMIOLÓGICA DA FEBRE AFTOSA PARA O ESTADO DE MATO GROSSO. Pesquisa em Andamento. Dados já tabulados.

As entrevistas foram transcritas em arquivo texto que será utilizado para análise de sentimento. A análise de sentimento, é uma técnica de processamento de linguagem natural que vem sendo utilizada em diversas áreas para avaliar o sentimento das pessoas, baseado na análise de textos escritos ou falas transcritas, incluindo-se a área de saúde (ALMOODI et al., 2020; JI, XUN, GELLER, 2013).

A análise de sentimento será realizada utilizando linguagem Python, além de um modelo de representação de linguagem VADER, treinada para a língua inglesa.

#### 1.1 Seleção de perguntas dos questionários

A metódica seleção de perguntas nos questionários da pesquisa foi fundamentada na imperativa necessidade de abordar de modo eficiente os quatro objetivos essenciais do

estudo (anexo). Uma diversidade de fatores exerce influência sobre a inclinação dos produtores, revendedores de produtos veterinários e médicos veterinários a notificar casos suspeitos, conferindo ao engajamento desses atores na vigilância de enfermidades um desafio de substancial relevância a ser transposto no âmbito do domínio veterinário. Com o intuito de potencializar essa inclinação, é de natureza primordial que se compreendam as percepções desses atores quanto aos riscos inerentes à eclosão de doença e ao papel das autoridades no âmbito veterinário.

Houveram perguntas específicas para identificar as percepções de produtores, servidores do INDEA, revendedores de produtos agropecuários e veterinários do setor privado em relação ao sistema de vigilância epidemiológica da febre aftosa no Mato Grosso. A fim de compreender a percepção desses atores, com intuito de desenvolver estratégias educativas e intervenções que aumentem o engajamento das partes no sistema de vigilância.

Portanto, a seleção limitada de perguntas foi estrategicamente realizada para otimizar a coleta de informações essenciais que se alinham aos objetivos do estudo. No apêndice, é apresentada uma descrição detalhada da razão pela qual cada uma das perguntas foi selecionada.

#### 1.2 Elaboração e estruturação de planilhas: organização eficiente de dados

##### 1.2.1 Dados do experimento

Previamente ao meu experimento foram realizadas entrevistas (que posteriormente foram transcritas) abrangendo um total de 42 produtores em diversos municípios no Estado do Mato Grosso - Brasil (Brasnorte, Cuiabá, Juara, Juína, Marcelândia, Novo Horizonte e Santo Antônio. Além disso, foram conduzidas entrevistas com 13 revendedoras de produtos veterinários, provenientes de Alto da Boa Vista, Brasnorte, Canarana, Curvelândia, Guaraná do Norte, Guarantã do Norte, Juara, Juína, Paranaíta, Primavera do Leste, Santa Rita do Trivelato, Sapazel, Terra Nova); 18 médicos veterinários (dos municípios de Brasnorte, Cáceres, Chapada dos Guimarães, Cuiabá, Gaúcha do Norte, Juara, Lucas do Rio Verde,

Nobres, Poxoréu, Primavera do Leste, Quatro Marcos, Ribeirão Castanheira e Santa Rita do Trivelato); e 19 servidores do INDEA-MT (não foram identificados os municípios onde desempenhavam suas funções).

#### 1.2.2 Organização dos dados

A etapa de organização dos dados implicou na elaboração de planilhas no formato Excel, sendo designada uma planilha específica para cada uma das categorias: produtores de gado, revendas veterinárias, médicos veterinários particulares e servidores do INDEA-MT. As indagações selecionadas para cada categoria foram dispostas na primeira coluna das respectivas planilhas. Simultaneamente, os nomes dos municípios foram inseridos na primeira linha, conferindo identificação às propriedades rurais dos entrevistados, bem como aos locais das revendas e aos campos de atuação dos médicos veterinários. No que concerne aos servidores do INDEA, estes foram designados numericamente de 1 a 19, correspondendo à ordem sequencial em que foram apresentados nas transcrições. O procedimento de organização dos dados almejou a disposição metódica e metódica das informações coletadas em cada categoria.

#### 1.2.3 Edição das transcrições

As respostas dos entrevistados, que previamente haviam sido transcritas utilizando a linguagem natural, foram subsequentemente incorporadas às células que correspondiam às perguntas pertinentes. Esse processo resultou na criação de um registro abrangente e estruturado das informações coletadas durante as entrevistas. Após a finalização da etapa de organização, foi produzida uma duplicata da planilha com o propósito de permitir a edição textual e sua subsequente tradução para o idioma da língua inglesa.

A linguagem natural representa o método de comunicação inerente aos seres humanos no decorrer de suas atividades diárias, com a finalidade de expressão e interação. Constitui o meio pelo qual os indivíduos estabelecem diálogos tanto de forma oral quanto escrita, empregando vocabulário, estruturas frasais e expressões usuais. Caracteriza-se por sua complexidade e diversidade, uma vez que engloba todas as sutilezas, ambiguidades e

variações que são intrínsecas à comunicação humana. A linguagem natural assume um papel fundamental na comunicação interpessoal e desempenha um papel central em várias tecnologias de processamento de linguagem natural, incluindo chatbots, assistentes de voz e sistemas de tradução automática (MARTIN, 2008). Por tanto se fez necessário a edição desses textos a qual refere-se à revisão e aprimoramento do texto transcrito para maior clareza e coesão. Esse processo é fundamental para preservar e tornar acessíveis informações históricas e relatos pessoais. O resultado é um registro escrito fiel às entrevistas orais, facilitando a pesquisa, análise e compartilhamento dessas narrativas (BAUM, 2008).

A edição das transcrições de histórias orais é um processo importante para garantir a precisão, clareza e fidelidade ao conteúdo original das entrevistas. A edição cuidadosa é necessária para corrigir esses problemas e produzir um registro fiel e legível das informações compartilhadas durante a entrevista (BAUM 2008). Durante a fase de edição dos textos, foi essencial realizar correções em relação a erros gramaticais, bem como garantir a coesão e coerência dos conteúdos. Essas intervenções eram totalmente apropriadas, considerando que as transcrições foram mantidas fielmente ao modo como os entrevistados se expressaram. Além disso, foram removidas expressões características de conversas informais do dia a dia, frequentemente envolvendo gírias, abreviações e variações linguísticas que não são tão usuais na linguagem formal. Empregou-se um considerável empenho em efetuar alterações mínimas no texto, visando à preservação do sentido original, com o propósito de assegurar uma análise mais precisa e fiel.

#### 1.2.4 Tradução dos textos editados para o inglês

Após a conclusão da edição dos textos, procedeu-se à tradução para o inglês com o auxílio do DeepL ([deepl.com/translator](https://www.deepl.com/translator)), um serviço gratuito de tradução que opera por meio de redes neurais, semelhante ao Google Tradutor. Essa ferramenta possibilita a

tradução do inglês para o português, no entanto, sua tecnologia vai além, ao proporcionar conversões de idioma com um aspecto mais natural, sobretudo em sentenças longas e complexas (HIDALGO-TERNERO, 2020).

1.2.5 Incorporação das respostas conforme as perguntas realizadas de forma agrupada pelo entrevistador para maior preservação de dados. Durante a etapa de composição das planilhas e revisão dos textos, deparou-se com algumas complexidades. Os entrevistadores não seguiram rigidamente o roteiro de perguntas, o que resultou na omissão de várias indagações e, por conseguinte, na existência de células vazias. Esse padrão se manifestou de maneira notória nas planilhas relativas aos servidores do INDEA, revendedores e médicos veterinários, enquanto os entrevistadores dos produtores rurais se mostraram mais aderentes ao roteiro. Além do que, em determinados cenários, ocorreram situações nas quais as perguntas foram formuladas, mas os entrevistados optaram por não responder, escolhendo permanecer em silêncio perante a questão, algo que foi pouco frequente.

Outra observação de relevo é que algumas respostas fornecidas pelos entrevistados não se alinharam com as perguntas propostas, enquanto em outros casos as respostas se revelaram confusas e destituídas de coerência; esses padrões se mostraram mais frequentes nas entrevistas com produtores rurais.

Adicionalmente, em algumas instâncias, as respostas oferecidas foram notavelmente concisas, o que levou à decisão de agregar perguntas, como por exemplo as perguntas “Quando eu falo INDEA-MT o que vem à sua cabeça? Qual é a importância do INDEA-MT em Mato Grosso? E para você qual é a importância do INDEA na sua atividade pecuária? Como percebe a situação do INDEA-MT nos dias de hoje? Você sabe qual é a missão do INDEA-MT? Em sua opinião, o INDEA-MT está cumprindo sua missão? Por quê?”, que no roteiro original eram separadas, a fim de se ampliar o conteúdo para a posterior análise de dados mediante a ferramenta VADER, para que assim com um maior léxico, possa ser feita uma análise mais consistente.

Ademais, em certas circunstâncias, as respostas apresentadas demonstraram ser notavelmente concisas, suscitando a decisão de unir questões que originalmente eram tratadas de maneira independente no roteiro, a exemplo das seguintes perguntas: “Quando eu falo INDEA-MT o que vem à sua cabeça? Qual é a importância do INDEA-MT em Mato Grosso? E para você qual é a importância do INDEA na sua atividade pecuária? Como percebe a situação do INDEA-MT nos dias de hoje? Você sabe qual é a missão do INDEA-MT? Em sua opinião, o INDEA-MT está cumprindo sua missão? Por quê?”. Tal decisão foi tomada com a finalidade de enriquecer o conteúdo direcionado à subsequente análise de dados por meio da ferramenta VADER. O propósito subjacente é facultar um léxico de maior amplitude, o que por conseguinte viabiliza uma avaliação mais substancial e coerente.

## Resultados

A seleção das perguntas nos questionários foi cuidadosamente feita para abordar os quatro principais objetivos do estudo: entender as percepções de produtores, revendedores de produtos veterinários, médicos veterinários e servidores do INDEA-MT em relação à vigilância da febre aftosa no Mato Grosso. A organização dos dados envolveu a criação de planilhas específicas para cada categoria de entrevistados e a edição das transcrições das entrevistas para maior clareza e coesão. As respostas foram posteriormente traduzidas para o inglês com o auxílio do DeepL, e foi feito um esforço para agrupar as respostas de forma coerente às perguntas feitas pelos entrevistadores.

Durante o processo, surgiram algumas complexidades, como a omissão de perguntas, células vazias, respostas que não se alinhavam com as perguntas e respostas concisas. Para superar essas complexidades, algumas perguntas foram agrupadas para enriquecer o léxico e facilitar a análise de dados usando a ferramenta VADER.

Em resumo, a seleção de perguntas, organização e edição dos dados visaram tornar

a coleta e análise de informações mais eficazes, considerando a variedade de respostas e complexidades encontradas durante as entrevistas.

## Conclusões

A conclusão do processo de seleção de perguntas, organização de dados e edição de transcrições foi a criação de um conjunto de dados estruturado e coerente que abrangeu entrevistas com produtores rurais, revendedores de produtos veterinários, médicos veterinários e servidores do INDEA-MT. Esse conjunto de dados foi preparado para análise posterior, incluindo a aplicação da ferramenta VADER para avaliar o sentimento das respostas em relação à vigilância da febre aftosa no Mato Grosso.

Durante esse processo, várias complexidades foram abordadas, como a omissão de perguntas, células vazias e respostas que não se alinhavam com as perguntas originais. Além disso, algumas perguntas foram agrupadas para aumentar o léxico disponível para análise. No geral, a conclusão é que a preparação dos dados foi fundamental para tornar o material coletado mais adequado para análise, permitindo uma compreensão mais profunda das percepções e opiniões dos entrevistados em relação à vigilância da febre aftosa no Mato Grosso. A análise subsequente desses dados será essencial para atingir os objetivos do estudo e fornecer insights valiosos sobre como melhorar o engajamento das partes interessadas na vigilância de doenças veterinárias.

## Agradecimentos

Gostaria de expressar minha profunda gratidão ao Professor Fernando Ferreira, meu orientador, cuja orientação, gentileza e paciência foram fundamentais para o sucesso deste trabalho. Sua dedicação à minha formação acadêmica foi inestimável.

## Referências

- ALAMOUDI, A. H.; ZAIDAN, B. B.; ZAIDAN, A. A.; ALBAHRI, O. S.; MOHAMMED, K. I.; MALIK, R. Q.; ALMAHDI, E. M.; CHYAD, M. A.; TAREQ, Z.; ALBAHRI, A. S.; HAMEED, H.; ALAA, M. **Sentiment analysis and its applications in fighting covid-19 and infectious diseases: a systematic review.** *Expert Systems with Applications*, v. 167, e. 114155, 2021.
- BARROS, Lucas; TRIFAN, Alina; OLIVEIRA, José Luís. **VADER meets BERT: sentiment analysis for early detection of signs of self-harm through social mining.** In: CLEF (Working Notes). 2021. p. 897-907.
- BAUM, Willa K. **Transcribing and Editing Oral History.** 5. ed. Association for State and Local History, 2008.
- BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Secretaria de Defesa Agropecuária. Departamento de Saúde Animal. **Programa Nacional de Erradicação e Prevenção da Febre Aftosa (PNEFA): plano estratégico 2017 – 2026, atualização 2019.** Brasília: MAPA, 2019. Disponível em: <http://www.indea.mt.gov.br/documents/363967/8520184/Plano+Estrat%C3%A9gico+PNEFA+2017-2026+-+Atualiza%C3%A7%C3%A3o+2019.pdf/564d5f31-63a3-c38b-12df-816c3f6d1fe8>. Acesso em: 6 de julho de 2023.
- ELBAGIR, Shihab; YANG, Jing. **Twitter sentiment analysis using natural language toolkit and VADER sentiment.** In: Proceedings of the international multiconference of engineers and computer scientists. 2019. p. 16.
- HIDALGO-TERNERO, Carlos Manuel. **Google Translate vs. DeepL. MonTI.** Monografias de Traducción e Interpretación, v. 12, n. 12, p. 154-177, 2020.
- JI, X.; CHUN, S. A.; GELLER, J. **Monitoring public health concerns using twitter sentiment classifications.** In: IEEE INTERNATIONAL CONFERENCE ON HEALTHCARE INFORMATICS, ICHI 2013, Philadelphia. Proceedings [...]. Philadelphia, 2013. p. 335–344.
- KOTHARI, Anita et al. **How do Canadian public health agencies respond to the COVID-19 emergency using social media: a**



**protocol for a case study using content and sentiment analysis.** BMJ open, v. 11, n. 4, p. e041818, 2021.

MARTIN, James H.; JURAFSKY, Dan. **Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition.** 2ª ed. Prentice Hall, 30 de março de 2008.

PAHO – PANAMERICAN HEALTH ORGANIZATION. **Hemispheric Program for the Eradication of Foot-and-Mouth Disease – PHEFA**, Beyond 2020. Technical Note. Cartagena, 2019.

SEABOLD, S.; PERKTOLD, J. Statsmodels: **Econometric and statistical modeling with python.** In: PYTHON IN SCIENCE CONFERENCE, 9., 2010. Proceedings [...]. VEENA, G.; VINAYAK, Aadithya; NAIR, Anu J. **Sentiment analysis using improved VADER and dependency parsing.** In: 2021 2nd Global Conference for Advancement in Technology (GCAT). IEEE, 2021. p. 1-6.

# **EVALUATION OF THE FEELINGS OF PRODUCERS AND ANIMAL HEALTH PROFESSIONALS IN THE STATE OF MATO GROSSO REGARDING THE EPIDEMIOLOGICAL SURVEILLANCE SYSTEM FOR FOOT-AND-MOUTH DISEASE**

**Helena Martos Romboli**

**Fernando Ferreira**

UNIVERSIDADE DE SÃO PAULO  
FACULDADE DE MEDICINA VETERINÁRIA E ZOOTECNIA (FMVZ/USP)

helenamartosromboli@usp.br

## **Objectives**

Use artificial intelligence techniques to assess the feelings of producers, INDEA officials, producers' resellers, INDEA officials, agricultural product resellers and private service veterinarians regarding different aspects of the epidemiological surveillance system for foot-and-mouth disease in the state of Mato Grosso. To verify the applicability of these tools to promote the engagement of these actors in the FMD surveillance system.

## **Materials and Methods**

The Laboratory of Epidemiology and Biostatistics (LEB), of the Department of Preventive Veterinary Medicine and Single Animal Health (VPS) at FMVZ/USP, has been working on evaluating the surveillance system for FMD in the state of Mato Grosso. The surveys were carried out between 09/11/2020 and 02/15/2021 in various municipalities in the state of Mato Grosso.

In order to carry out the surveys, the project was presented to the State Management Team of the 2017-2026 Strategic Plan of the National Program for the Eradication and Prevention of Foot-and-Mouth Disease, within the scope of the Institute of Agricultural Defense of the State

of Mato Grosso, which is made up of public and private organizations (MATO GROSSO, 2018), whose support made it possible to hire the outsourced company with the competence and experience required for the project, Vetor Pesquisa, from the Emergency Animal Health Fund, in order to avoid the bias of the interviewer belonging to the SVO. The company took part in the planning phase, drawing up the questionnaires and scripts, carrying out the field research and drawing up reports with exploratory analysis of the data and perceptions from the qualitative interviews (VETOR, 2021). At the end of the work, it provided the databases, which are used for analysis.

The scripts for the in-depth interviews were designed to capture perceptions that would not be possible through the application of the structured questionnaire alone, such as the factors that can influence the disease notification process and the beliefs and behaviors related to the surveillance system.

The following topics were covered when constructing the scripts: sociodemographic information, evaluation of the information systems provided by INDEA, adoption of biosecurity measures in rural establishments, knowledge of foot-and-mouth disease, perceptions of disease notification and the stages involved, suspension of vaccination against foot-and-mouth disease and the

institutional image of INDEA and the Emergency Animal Health Fund (FESA - MT). Criteria were established for selecting the public to take part in the interviews: rural producers or those responsible for the management of rural establishments with cattle; self-employed veterinarians who work with production animals, and retailers selling drugs and biologicals for animal production.

Once the interview scripts had been finalized, they were tested in the field by conducting pilot interviews, where they checked that the questions were easy to understand, especially for people with low levels of education, that the answers were not being collected in a biased way and that the order of the questions was correct so as not to induce responses.

The teams responsible for the interviews were trained by the company in advance.

Table 2 shows the sample description for each population and the sample size obtained. As this is qualitative research, the process of which is more complex and time-consuming, the sample size was estimated taking into account the collection capacity during the project period.

Tabela 2: Descrição das amostras de pesquisa, considerando o público-alvo, universo, amostra e qualidade. Mato Grosso, 2020.

Público Alvo	Universo	Amostra qualitativa
Servidores do INDEA	821	20
Produtores rurais	108.079	40
Médico veterinário	1.255	20
Revendas de produtos veterinários	1.006	20

Fonte: SCHMIDT, Ana Carolina. AVALIAÇÃO E PROPOSTA DE REESTRUTURAÇÃO DO SISTEMA DE VIGILÂNCIA EPIDEMIOLÓGICA DA FEBRE AFTOSA PARA O ESTADO DE MATO GROSSO. Pesquisa em Andamento. Dados já tabulados.

The interviews were transcribed into a text file which will be used for sentiment analysis. Sentiment analysis is a natural language processing technique that has been used in several areas to assess people's feelings, based on the analysis of written texts or transcribed speech, including the health area (ALMOODI et al., 2020; JI, XUN, GELLER, 2013).

Sentiment analysis will be carried out using the Python language, as well as a VADER language representation model, trained for the English language.

#### 1.1 Selection of survey questions

The meticulous selection of questions in the survey questionnaires was based on the imperative need to efficiently address the four essential objectives of the study (appendix). A

variety of factors influence the inclination of producers, retailers of veterinary products and veterinarians to report suspected cases, making the engagement of these actors in disease surveillance a challenge of substantial relevance to be overcome in the veterinary field. In order to boost this inclination, it is essential to understand the perceptions of these actors regarding the risks inherent in the outbreak of disease and the role of the authorities in the veterinary field.

There were specific questions to identify the perceptions of producers, INDEA officials, agricultural product retailers and private sector veterinarians in relation to the epidemiological surveillance system for foot-and-mouth disease in Mato Grosso. The aim was to understand the perceptions of these actors, with a view to developing educational strategies and interventions to increase their engagement in the surveillance system.

Therefore, the limited selection of questions was strategically carried out to optimize the collection of essential information that aligns with the objectives of the study. A detailed description of why each of the questions was selected is provided in the appendix.

#### 1.2 Spreadsheet design and structuring: efficient data organization

##### 1.2.1 Data from the experiment

Prior to my experiment, interviews were conducted (which were later transcribed) with a total of 42 producers in various municipalities in the state of Mato Grosso - Brazil (Brasnorte, Cuiabá, Juara, Juína, Marcelândia, Novo Horizonte and Santo Antônio. In addition, interviews were conducted with 13 veterinary product retailers from Alto da Boa Vista, Brasnorte, Canarana, Curvelândia, Guaraná do Norte, Guarantã do Norte, Juara, Juína, Paranaíta, Primavera do Leste, Santa Rita do Trivelato, Sapazel, Terra Nova); 18 veterinarians (from the municipalities of Brasnorte, Cáceres, Chapada dos Guimarães, Cuiabá, Gaúcha do Norte, Juara, Lucas do Rio Verde, Nobres, Poxoréu, Primavera do Leste, Quatro Marcos, Ribeirão Castanheira and Santa Rita do Trivelato); and 19 INDEA-MT



employees (the municipalities where they performed their duties were not identified).

#### 1.2.2 Organizing the data

The data organization stage involved preparing spreadsheets in Excel format, with a specific spreadsheet being assigned to each of the categories: cattle producers, veterinary dealers, private veterinarians and INDEA-MT employees. The questions selected for each category were placed in the first column of the respective spreadsheets. At the same time, the names of the municipalities were inserted in the first line, identifying the rural properties of the interviewees, as well as the locations of the dealers and the fields of work of the veterinarians. The INDEA employees were numbered from 1 to 19, corresponding to the sequential order in which they were presented in the transcripts. The data organization procedure aimed to provide a meticulous and methodical arrangement of the information collected in each category.

#### 1.2.3 Editing the transcripts

The interviewees' answers, which had previously been transcribed using natural language, were subsequently incorporated into the cells corresponding to the relevant questions. This process resulted in the creation of a comprehensive and structured record of the information collected during the interviews. Once the organization stage was complete, a duplicate of the spreadsheet was produced to allow for textual editing and subsequent translation into English.

Natural language represents the method of communication inherent to human beings in the course of their daily activities, with the aim of expression and interaction. It is the means by which individuals establish dialogues both orally and in writing, using vocabulary, sentence structures and common expressions. It is characterized by its complexity and diversity, since it encompasses all the subtleties, ambiguities and variations that are intrinsic to human communication. Natural language plays a fundamental role in interpersonal communication and plays a central role in various natural language processing technologies, including chatbots, voice

assistants and machine translation systems (MARTIN, 2008). It was therefore necessary to edit these texts, which refers to the revision and improvement of the transcribed text for greater clarity and cohesion. This process is fundamental to preserving and making accessible historical information and personal accounts. The result is a written record that is faithful to the oral interviews, making it easier to research, analyze and share these narratives (BAUM, 2008).

Editing oral history transcripts is an important process to ensure accuracy, clarity and fidelity to the original content of the interviews. Careful editing is necessary to correct these problems and produce a faithful and legible record of the information shared during the interview (BAUM 2008). During the editing phase of the texts, it was essential to make corrections in relation to grammatical errors, as well as ensuring the cohesion and coherence of the content. These interventions were entirely appropriate, considering that the transcripts were kept faithful to the way the interviewees expressed themselves. In addition, expressions characteristic of everyday informal conversations were removed, often involving slang, abbreviations and linguistic variations that are not so usual in formal language. Considerable effort was put into making minimal changes to the text in order to preserve the original meaning, with the aim of ensuring a more accurate and faithful analysis.

#### 1.2.4 Translating the edited texts into English

Once the texts had been edited, they were translated into English using DeepL ([deepl.com/translator](https://www.deepl.com/translator)), a free translation service that uses neural networks, similar to Google Translate. This tool makes it possible to translate from English into Portuguese, but its technology goes further by providing more natural-looking language conversions, especially for long, complex sentences (HIDALGO-TERNERO, 2020).

#### 1.2.5 Incorporation of answers according to the questions asked in a grouped manner by the interviewer for greater data preservation

During the stage of composing the spreadsheets and revising the texts, some

complexities were encountered. The interviewers didn't strictly follow the script of questions, which resulted in the omission of several questions and, consequently, the existence of empty cells. This pattern was notorious in the spreadsheets for INDEA employees, retailers and veterinarians, while the interviewers for rural producers were more adherent to the script. Furthermore, in certain scenarios, there were situations in which the questions were asked, but the interviewees chose not to answer, choosing to remain silent on the issue, something that was infrequent. Another noteworthy observation is that some of the answers provided by the interviewees were not in line with the proposed questions, while in other cases the answers were confusing and lacked coherence; these patterns were more frequent in the interviews with rural producers. Additionally, in some instances, the answers offered were remarkably concise, which led to the decision to add questions, such as "When I say INDEA-MT, what comes to mind? What is the importance of INDEA-MT in Mato Grosso? And what do you think is the importance of INDEA in your livestock activity? How do you perceive INDEA-MT's situation today? Do you know what INDEA-MT's mission is? In your opinion, is INDEA-MT fulfilling its mission? Why?", which in the original script were separate, in order to broaden the content for later data analysis using the VADER tool, so that with a larger lexicon, a more consistent analysis could be made. Furthermore, in certain circumstances, the answers provided proved to be remarkably concise, prompting the decision to merge questions that were originally dealt with independently in the script, such as the following questions: "When I say INDEA-MT, what comes to mind? What is the importance of INDEA-MT in Mato Grosso? And what do you think is the importance of INDEA in your livestock activity? How do you perceive INDEA-MT's situation today? Do you know what INDEA-MT's mission is? In your opinion, is INDEA-MT fulfilling its mission? Why?". This decision was made in order to enrich the content for subsequent data analysis using the

VADER tool. The underlying purpose is to provide a broader lexicon, which therefore enables a more substantial and coherent evaluation.

## Results

The questions in the questionnaires were carefully selected to address the four main objectives of the study: to understand the perceptions of producers, veterinary product retailers, veterinary doctors and INDEA-MT officials regarding foot-and-mouth disease surveillance in Mato Grosso. Organizing the data involved creating specific spreadsheets for each category of interviewee and editing the interview transcripts for clarity and cohesion. The answers were then translated into English with the help of DeepL, and an effort was made to group the answers in a way that was coherent with the questions asked by the interviewers.

During the process, some complexities arose, such as the omission of questions, empty cells, answers that did not align with the questions and concise answers. To overcome these complexities, some questions were grouped together to enrich the lexicon and facilitate data analysis using the VADER tool.

In summary, the selection of questions, organization and editing of the data aimed to make the collection and analysis of information more effective, considering the variety of answers and complexities encountered during the interviews.

## Conclusions

The conclusion of the process of selecting questions, organizing data and editing transcripts was the creation of a structured and coherent data set covering interviews with rural producers, veterinary product retailers, veterinary doctors and INDEA-MT officials. This dataset was prepared for further analysis, including the application of the VADER tool to assess the sentiment of the responses in relation to foot-and-mouth disease surveillance in Mato Grosso.

During this process, various complexities were addressed, such as the omission of questions, empty cells and answers that did not align with the original questions. In addition, some questions were grouped together to increase the lexicon available for analysis.

Overall, the conclusion is that data preparation was fundamental in making the material collected more suitable for analysis, allowing for a deeper understanding of the interviewees' perceptions and opinions regarding foot-and-mouth disease surveillance in Mato Grosso. Subsequent analysis of this data will be essential to achieve the study's objectives and provide valuable insights into how to improve stakeholder engagement in veterinary disease surveillance.

### Acknowledgements

I would like to express my deepest gratitude to Professor Fernando Ferreira, my advisor, whose guidance, kindness and patience were fundamental to the success of this work. His dedication to my academic training has been invaluable.

### References

ALAMOUDI, A. H.; ZAIDAN, B. B.; ZAIDAN, A. A.; ALBAHRI, O. S.; MOHAMMED, K. I.; MALIK, R. Q.; ALMAHDI, E. M.; CHYAD, M. A.; TAREQ, Z.; ALBAHRI, A. S.; HAMEED, H.; ALAA, M. **Sentiment analysis and its applications in fighting covid-19 and infectious diseases: a systematic review.** *Expert Systems with Applications*, v. 167, e. 114155, 2021.

BARROS, Lucas; TRIFAN, Alina; OLIVEIRA, José Luís. **VADER meets BERT: sentiment analysis for early detection of signs of self-harm through social mining.** In: CLEF (Working Notes). 2021. p. 897-907.

BAUM, Willa K. **Transcribing and Editing Oral History.** 5. ed. Association for State and Local History, 2008.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Secretaria de Defesa Agropecuária. Departamento de Saúde Animal. **Programa Nacional de Erradicação e Prevenção da Febre Aftosa (PNEFA): plano**

**estratégico 2017 – 2026, atualização 2019.** Brasília: MAPA, 2019. Disponível em: <http://www.indea.mt.gov.br/documents/363967/8520184/Plano+Estrat%C3%A9gico+PNEFA+2017-2026+-+Atualiza%C3%A7%C3%A3o+2019.pdf/564d5f31-63a3-c38b-12df-816c3f6d1fe8> . Acesso em: 6 de julho de 2023.

ELBAGIR, Shihab; YANG, Jing. **Twitter sentiment analysis using natural language toolkit and VADER sentiment.** In: Proceedings of the international multiconference of engineers and computer scientists. 2019. p. 16.

HIDALGO-TERNERO, Carlos Manuel. **Google Translate vs. DeepL. MonTI.** Monografías de Traducción e Interpretación, v. 12, n. 12, p. 154-177, 2020.

JI, X.; CHUN, S. A.; GELLER, J. **Monitoring public health concerns using twitter sentiment classifications.** In: IEEE INTERNATIONAL CONFERENCE ON HEALTHCARE INFORMATICS, ICHI 2013, Philadelphia. Proceedings [...]. Philadelphia, 2013. p. 335–344.

KOTHARI, Anita et al. **How do Canadian public health agencies respond to the COVID-19 emergency using social media: a protocol for a case study using content and sentiment analysis.** *BMJ open*, v. 11, n. 4, p. e041818, 2021.

MARTIN, James H.; JURAFSKY, Dan. **Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition.** 2ª ed. Prentice Hall, 30 de março de 2008.

PAHO – PANAMERICAN HEALTH ORGANIZATION. **Hemispheric Program for the Eradication of Foot-and-Mouth Disease – PHEFA, Beyond 2020.** Technical Note. Cartagena, 2019.

SEABOLD, S.; PERKTOLD, J. **Statsmodels: Econometric and statistical modeling with python.** In: PYTHON IN SCIENCE CONFERENCE, 9., 2010. Proceedings [...].

VEENA, G.; VINAYAK, Aadithya; NAIR, Anu J. **Sentiment analysis using improved VADER and dependency parsing.** In: 2021 2nd Global



Conference for Advancement in Technology  
(GCAT). IEEE, 2021. p. 1-6.