

Poultry in Motion: Text Mining Social Media Data for Avian Influenza Surveillance in the UK

Authors : Samuel Munaf, Kevin Swingler, Franz Brülisauer, Anthony O'Hare, George Gunn, Aaron Reeves

Abstract : Background: Avian influenza, more commonly known as Bird flu, is a viral zoonotic respiratory disease stemming from various species of poultry, including pets and migratory birds. Researchers have purported that the accessibility of health information online, in addition to the low-cost data collection methods the internet provides, has revolutionized the methods in which epidemiological and disease surveillance data is utilized. This paper examines the feasibility of using internet data sources, such as Twitter and livestock forums, for the early detection of the avian flu outbreak, through the use of text mining algorithms and social network analysis. Methods: Social media mining was conducted on Twitter between the period of 01/01/2021 to 31/12/2021 via the Twitter API in Python. The results were filtered firstly by hashtags (#avianflu, #birdflu), word occurrences (avian flu, bird flu, H5N1), and then refined further by location to include only those results from within the UK. Analysis was conducted on this text in a time-series manner to determine keyword frequencies and topic modeling to uncover insights in the text prior to a confirmed outbreak. Further analysis was performed by examining clinical signs (e.g., swollen head, blue comb, dullness) within the time series prior to the confirmed avian flu outbreak by the Animal and Plant Health Agency (APHA). Results: The increased search results in Google and avian flu-related tweets showed a correlation in time with the confirmed cases. Topic modeling uncovered clusters of word occurrences relating to livestock biosecurity, disposal of dead birds, and prevention measures. Conclusions: Text mining social media data can prove to be useful in relation to analysing discussed topics for epidemiological surveillance purposes, especially given the lack of applied research in the veterinary domain. The small sample size of tweets for certain weekly time periods makes it difficult to provide statistically plausible results, in addition to a great amount of textual noise in the data.

Keywords : veterinary epidemiology, disease surveillance, infodemiology, infoveillance, avian influenza, social media

Conference Title : ICSMMSNA 2022 : International Conference on Social Media Mining and Social Network Analysis

Conference Location : Vancouver, Canada

Conference Dates : May 23-24, 2022