MUST-AI: Multisource Surveillance Tool -Avian Influenza

Carlene Trevennec, Pierre Pompidor, Samira Bououda, Julien Rabatel, Mathieu Roche

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2024) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

28th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2024)

MUST-AI: Multisource Surveillance Tool - Avian Influenza

Carlène Trevennec[a], Pierre Pompidor[b], Samira Bououda[c], Julien Rabatel[d], Mathieu Roche[e,*]

[a]*INRAE, Montpellier, France / ASTRE, Univ Montpellier, CIRAD, INRAE, Montpellier, France*
[b]*LIRMM, Université Montpellier, CNRS, Montpellier, France*
[c]*ASTRE, Univ Montpellier, CIRAD, INRAE, Montpellier, France*
[d]*CIRAD, F-34398 Montpellier*
[e]*CIRAD, F-34398 Montpellier, France / TETIS, Univ Montpellier, AgroParisTech, CIRAD, CNRS, INRAE, Montpellier, France*

Abstract

The multisource surveillance tool (MUST) is a platform for collecting, gathering, and visualizing different sources of information related to health events and highly pathogenic avian influenza in mammals (HPAIM). MUST-AI constitutes the first part of the MUST tool, which centralizes health information relating to cases of HPAIM since January 1, 2021, and comes from 3 different notification sources, an official notification source confirmed by public health institutions (i.e., WAHIS) and two other alternative unofficial sources that collect events from online media (PADI-web) and expert networks (ProMED). Owing to the use of natural language processing (NLP) algorithms, HPAIM events are represented on an interactive map associated with a graph that represents their distribution over a given time interval. This paper presents new tools and approaches for data fusion and experiments for selecting data to integrate into MUST that are related to HPAIM events.

Keywords: Epidemic intelligence; Event-based surveillance; Data fusion; Highly pathogenic avian influenza

1. Introduction

Zoonotic infectious diseases pose a significant threat to public health, and it is essential to detect them early and respond promptly to minimize risk at the earliest possible stage. Traditional animal health surveillance systems involve a step-by-step process of field suspicion, laboratory confirmation and formal notification to international organizations, such as the World Organization for Animal Health (WOAH). This process ensures the transparency and accuracy of global animal health data [10].

* Corresponding author. Tel.: +33467558612
E-mail address: mathieu.roche@cirad.fr

Avian influenza (AI) is a highly contagious viral disease that affects both domestic and wild birds. This zoonotic virus has been the source of worldwide outbreaks, leading to several million wild bird deaths, culled poultry, and sporadic human cases. Recently, there has been an increase in reported infections in mammals, ranging in intensity from no symptoms to mass mortality events [6]. As a WOAH-listed disease, the occurrence of AI in unusual host species is considered an exceptional epidemiological event and thus must be noted through the World Animal Health Information System (WAHIS) within 24 hours after confirmation [10]. However, several examples from the scientific literature have documented AI evidence in mammals that has not been reported or identified within weeks or months [14, 1, 16]. Similarly, it has been observed that some disease outbreaks are only recorded in the national health information systems of specific countries. This can lead to discrepancies when comparing publicly available official national counts of outbreaks and WAHIS reports[1]. In some cases, underreporting of outbreaks can result in gaps in the information available, such as the geographical coverage and the range of affected species, and may result in delayed notifications. Consequently, the awareness of virus spread among unusual hosts might be incomplete and delayed.

In the last two decades, this recurrent challenge has led animal health agency teams to use complementary approaches and technologies to reinforce the capacity of the surveillance systems to detect emerging infectious diseases [13]. Web-based early warning systems, such as ProMED-mail, the Global Public Health Intelligence Network (GPHIN), Health Maps, and PADI-web, have been able to recognize emerging infectious diseases earlier than the traditional surveillance systems [7, 11, 15]. These systems are highly adaptable, low-cost and/or operate in real time, all of which are necessary features for the early detection of emerging diseases. They offer the opportunity to capture complementary information and improve surveillance coverage (e.g., spatially and temporally). At the same time, these methods produce large volumes of unstructured and non-validated data that can lead to difficult and time-consuming interpretations or inaccurate health status predictions [2, 7].

A new surveillance tool called the multisource surveillance tool (MUST[2]) has been developed. This paper presents an original platform for collecting, compiling, and visualizing different sources of information related to the relevant events and the status of HPAIM (highly pathogenic avian influenza in mammals). The first version of the MUST tool is called "MUST-AI", which filters and centralizes health information related to cases of HPAIM (excluding human cases) starting from January 1st, 2021. The tool collects information from three different notification sources, an official notification source confirmed by public health institutions (WAHIS), and two other unofficial alternative sources that collect events from online media (PADI-web data) and networks of experts (ProMED-mail). The MUST interface has been designed to ease risk assessments by providing (i) data visualizations to assess spatiotemporal coverage in real time, (ii) a list of affected species to adapt to surveillance programs and (iii) direct access to articles to obtain additional details about the detected events.

The rest of the paper is organized as follows. Section 2 outlines related work in animal disease surveillance based on WAHIS, ProMED, and PADI-web; Section 3 presents the fusion methods of these systems integrated into our original tool called MUST; Section 4 evaluates the type of events related to HPAIM extracted with PADI-web to integrate into MUST. Finally, future directions for this research are proposed in Section 5.

2. Processing of data sources

2.1. WAHIS

WAHIS[3] is the global animal health reference database of the World Organization for Animal Health (WOAH). WAHIS data reflect the validated information collected since 2005 reported by the Veterinary Services from Member and Non-Member Countries and Territories on terrestrial and aquatic listed diseases in domestic animals and wildlife, as well as on emerging diseases and zoonoses. WAHIS includes interactive mapping tools and dashboards to support data consultations, visualizations, and extractions of officially validated animal health data.

[1] Animal and Plant Health Agency. 2023. "Confirmed Findings of Influenza of Avian Origin in Non-Avian Wildlife". Bird flu (avian influenza): findings in nonavian wildlife. GOV.UK. 31 October 2023 - https://www.gov.uk/government/publications/bird-flu-avian-influenza-findings-in-non-avian-wildlife/confirmed-findings-of-influenza-of-avian-origin-in-non-avian-wildlife

[2] http://must-surveillance.com/

[3] https://wahis.woah.org/#/home

2.2. ProMED

The Program for Monitoring Emerging Diseases (ProMED) [3] is a global system for reporting and tracking outbreaks of infectious diseases. Since 1994, this program of the International Society for Infectious Diseases (ISID) has monitored outbreaks of infectious diseases affecting humans, animals, and even plants. In most countries in the world, those involved in health monitoring disseminate or relay alerts by e-mail (the information can then be accessed via the ProMED website: https://promedmail.org/). The advantage of this information network is that it is generally much more responsive than official channels, but the disadvantage is that the information is poorly structured (ProMED should eventually provide access to a database).

As part of the MUST project, we used ProMED's API every week to collect the posts related to cases of avian influenza affecting mammals (with the exception of humans). The challenge of analyzing emails is manifold, bearing in mind that information about a new outbreak is often associated with a recall of outbreaks that have occurred in the past at the regional or national level. Our automatic system must therefore identify in the header of an email its link with an avian flu epidemic (in particular by identifying the serotypes involved), then in the body of the email, the mammal species (excluding humans) affected by the disease, the date of the outbreak, its location and, if possible, the number of individuals affected and their environment (farm, wild). The example post in Figure 1 is an almost ideal example, particularly because it does not include chronological references to other epidemic outbreaks.

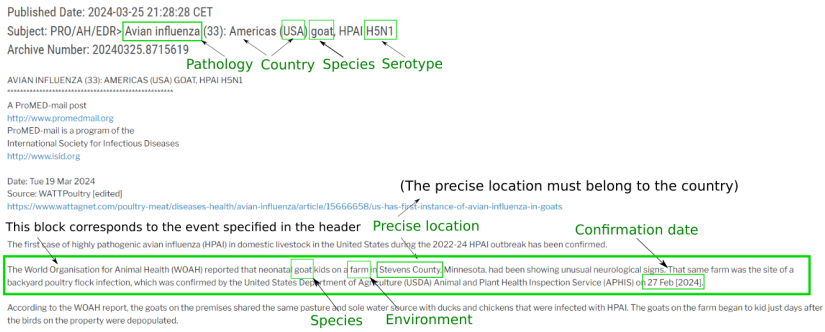

Fig. 1. Examples of elements extracted from a ProMED Post

This text analysis work is based on the following:

- A purpose-built lexicon of approximately 6,000 mammal species, including scientific names and vernacular names in English, Spanish and French.
- The GeoNames service, which provides detailed information on the place names; it performs an essential task by identifying the location of epidemic outbreaks.

Information extraction is based on regular expressions. While the analysis of relatively well-structured e-mail headers, in which the disease and its serotypes, the country (or continent) and the species of interest appear, is reliable, the analysis of the body of e-mails requires a check to verify the exact location of the epidemic (there is one inaccuracy per four posts). A final automated check is carried out to avoid duplication in the reporting of outbreaks (the main problem being that several e-mails may cite the same epidemic).

Note that evaluations based on 100 ProMED posts highlight that the scientific names of impacted mammal species are not given in 82% of posts. Moreover, the main problems identified are related to the extraction of irrelevant locations (i.e., 21% of locations are wrong or imprecise).

2.3. PADI-web

The Platform for Automated Extraction of Animal Disease Information from the Web (PADI-web) is an automated biosurveillance system devoted to online news source monitoring for the detection of emerging/new animal infectious

diseases via the French Epidemic Intelligence System. PADI-web[4] [15] automatically collects news via customized multilingual queries, classifies them and extracts epidemiological information. This tool was mainly developed for animal disease surveillance.

The PADI-web pipeline involves 5 steps ranging from online news collection to the extraction of epidemiological features. The 5 steps are: (1) data collection, (2) data processing, (3) data classification (i.e., news and sentence classification), (4) information extraction, and (5) visualization and user notification.

For Step 4, spaCy[5] was integrated into PADI-web 3.0 [15]. SpaCy already includes powerful named entity recognition (NER) models that allow the recognition of named entities. With PADI-web, we can use a classical model to identify well-known named entities such as locations and organizations. Moreover, specific information for the animal disease surveillance domain could be extracted by using specific dictionaries (e.g., host lexicons) and/or labeled datasets to learn and integrate a domain-specific model for NER. For location names, regular calls to the Geonames[6] gazetteer API aim to associate each recognized location name with a Geonames entity ID. Note that in the last version of the tool, five types of extractions are proposed to end-users:

- **Strategy A:** Events extracted from outbreak articles [spaCy locations in outbreak articles (i.e., document-based classification)];
- **Strategy B:** Events extracted from outbreak articles and current event sentences [spaCy locations in outbreak articles (i.e., document-based classification) and current event sentences (i.e., sentence-based classification)];
- **Strategy C:** Places extracted at the beginning of the outbreak articles [spaCy locations in the first 300 characters of the outbreak article];
- **Strategy D:** Events extracted in relevant articles based on PADI-web-specific locations [locations extracted with spaCy learned with labeled data];
- **Strategy E:** Events extracted at the beginning of relevant articles [spaCy locations in the first 300 characters of relevant articles].

The results of these strategies for extracting events to integrate into the MUST tool are discussed in Section 4.

3. Fusion of data with MUST

The multisource surveillance tool (MUST) tool is designed to provide a comprehensive view of various events. It consists of four main components, a map, some search criteria (above the map), a list of events (on the right of the map), and a distribution chart (below the map) (see Figure 2).

3.1. Filtering events

The filtering criteria enables the user to customize the events displayed on the map and other components based on their preferences. The user can specify the source of events, date range, and other parameters. After the criteria are modified, the displayed events are automatically updated on the fly.

One central aspect of events is their source, i.e., where they were first published. The user can filter the sources that should be displayed on the map by checking the corresponding checkboxes. The list of sources is dynamically generated from the events currently loaded in the GUI. Therefore, if a data source is not listed, it is not present in the currently loaded events. This can occur if the date range that is currently selected is narrow and does not contain events from all the sources. The colors are automatically assigned to each data source to visually differentiate them. The sources whose names start with ”Merged” have been created by MUST (see Figure 3). They correspond to the results of a process aimed at merging similar events (from one source or several distinct sources) into one event. The methods used for merging are detailed in Section 3.2. The name of these fusion results describe:

[4] https://www.padi-web-one-health.org
[5] https://spacy.io/
[6] https://www.geonames.org/

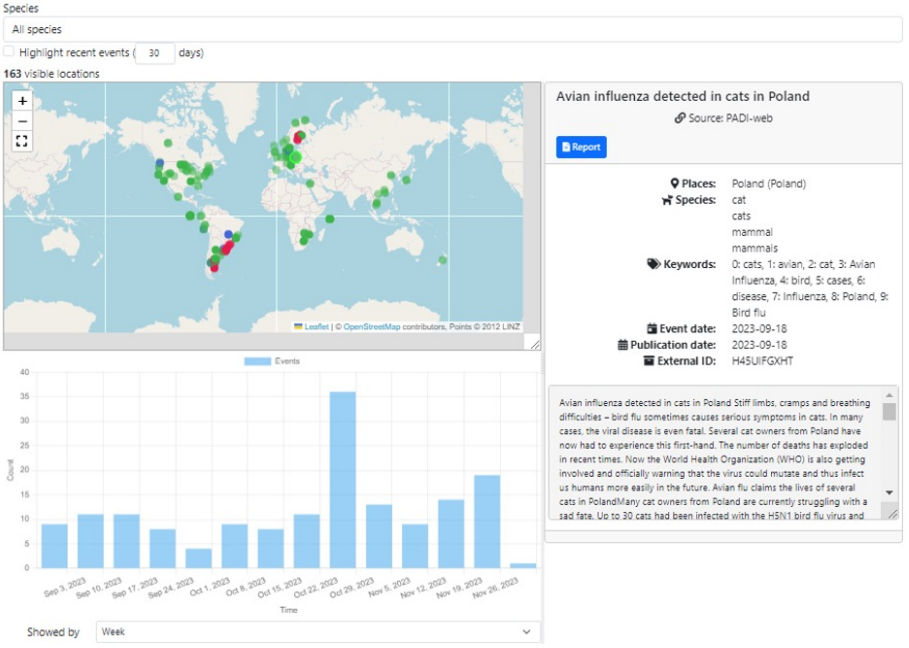

Fig. 2. General view of the MUST tool.

- The parameters used during the merging process, which are shown in parentheses.
- The sources that were merged to provide the resulting events. For instance, ”PADI-web + ProMED” contains all events that were obtained by merging events from PADI-web and ProMED.

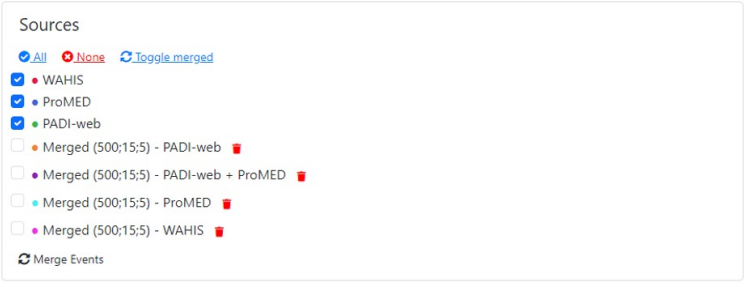

Fig. 3. Merging of events with the MUST tool.

The checkbox ”Highlight recent events” is not a filter per se but rather a way to visually highlight the most recent events on the map. All the event markers that have their event date within the last N days have a visible yellow circle around them (see the map in Figure 4).

3.2. Fusion method

This section describes the methodology and algorithms used to address end-user needs. Event merging is a tool designed to analyze and merge similar events based on certain criteria, even if they were published by different sources. The MUST tool works by comparing each event with every other event in the dataset. Three main criteria are used to determine whether two events are similar:

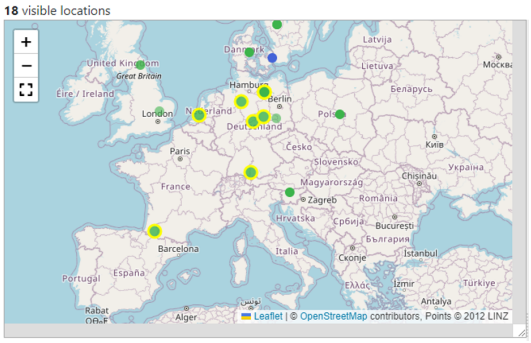

Fig. 4. Visualization of recent events.

1. **Geographical Proximity**: MUST calculates the geographical distance between events. The events are considered geographically close if this distance is less than a specified maximum distance.
2. **Temporal Proximity**: Our tool compares the dates of events. If the difference in dates is less than a specified maximum number of days, the events are considered temporally close.
3. **Species Similarity**: MUST compares the species involved in the events. This method uses a text comparison algorithm (i.e., the Levenshtein distance) to determine the similarity between the species names. The species are considered similar if the Levenshtein distance is less than a specified maximum value. If events have more than one species value, at least one pair of values must be fairly similar.

If two events meet all three criteria, they are considered similar and are grouped together. The application then merges the data of similar events, creating a new event that represents a group of similar events. MUST also provides a way to save merged events in a dedicated database, with each merged event being associated with a unique data source. The event merging algorithm follows these steps:

1. Preprocessing: The algorithm first converts all event data to a standard format.
2. Finding Similar Events: The algorithm then compares each event with every other event in the dataset using the three criteria mentioned above. It keeps track of which events are similar to each other.
3. Grouping Similar Events: Once all the comparisons are made, the algorithm groups together all the similar events. This is done using a depth-first search algorithm. For instance, if events A and B have been found to be similar and events B and C have also been found to be similar, the algorithm creates a group A, B, C.
4. Merging Similar Events: The algorithm then merges the data of similar events in each group, creating a new event that represents a group of similar events. The creation of the new merged event is proceeds as follows:
   • The first event of the group is used as a reference. It is cloned as a new event.
   • In addition to the event values coming from the reference event, we create new values coming from all the events from the group. The resulting merged event can have many event values for each data field (e.g., many locations, event dates).
   • The data source that is associated with the event is built as follows. Its name is ”Merged (500;15;5) - PADI-web + ProMED” (see Figure 3).
5. Saving Merged Events: Finally, the algorithm saves the merged events to the database.

4. Case study and results

4.1. Case study

It is mandatory to report the detection of the highly pathogenic avian influenza (HPAI) virus in an unusual host species to WAHIS. However, some countries either delay or fail to report such cases to the WAHIS database, which can jeopardize the quality of the early warning systems for these events.

Fig. 5. Visualization of the results obtained with MUST-AI in the USA. Red dot: WAHIS event / Green dot: PADI-web event

In parallel, several HPAIM findings have been documented in the literature. The European Food Safety Authority (EFSA) has published a global review of the literature on the role of mammals in avian influenza [6].

In this paper, we used the WAHIS database and some cases reported in published articles as our gold standard for identifying HPAIM cases. In total, we selected 7 case studies: 5 from the WAHIS database and 2 from the literature. These data correspond to 67 relevant events related to outbreaks associated with these case studies and were obtained with PADI-web.

The ease of associating a PADI-web article with a report or an official notification motivated the selection of certain case studies over others. More precisely, an event of which WAHIS was notified (i.e., red dots indicating reports with MUST) was selected for our study based on the number of WAHIS events represented in a given area. Conversely, in dedicated areas, there is a low concentration of WAHIS notifications, which facilitates analysis of the correlated PADI-web articles. This enables them to be integrated into our gold standard dataset.

Note that identifying a link between an officially reported outbreak of HPAIM and a PADI-web article can be difficult when certain areas have concentrations of both numerous WAHIS events and PADI-web articles, as is the case in the United States (see Figure 5), where PADI-web articles are difficult to associate with a particular WAHIS notification.

The scientific literature cannot cover recent HPAIM cases, as there is an average of one year from the detection of an outbreak to publication. Furthermore, some countries, such as China, India and Australia, were not selected as case studies because their cases were published too long ago (before the creation of PADI-web). The study could not consider the African continent because it suffers from a lack of surveillance resources, leading to underreporting on WAHIS and in the scientific literature.

In addition, this study did not target outbreaks of HPAIM due to very large quantities of PADI-web articles associated with these events, which complicates the evaluation described in Section 4.2. An example of this occurred in Spain during a farmed mink HPAI outbreak in October 2022. Approximately 83 PADI-web articles were reported, making it challenging to analyze the event [1].

4.2. Evaluation protocol and results

The selection of PADI-web articles for the construction of case studies is based on the similarity of the criteria date and the location of the outbreak mentioned in the article with the reference event reported to WAHIS or in the literature. Only PADI-web articles whose publication dates were between the date of the event and approximately one month after the date of the report or official notification were retained for this study. This time window was chosen based on previous work and could be refined in future work [15]. Articles for which at least one of these criteria was missing were not retained for this study; therefore, they constituted an expert prefilter in this selection process.

Once the study cases and associated PADI-web articles were identified, the aim was to evaluate the extraction quality of each strategy (see Section 2.3) in terms of recall and precision (formula (1)). Recall, also called sensitivity, corresponds to the number of relevant outbreaks found for each strategy out of the total number of relevant events actually present; in other words, recall makes it possible to determine whether all the relevant events that should have been reported by the strategies have been identified. Precision corresponds to the number of relevant outbreaks found

for each strategy out of the number of events returned by the automated system. If the system returned other outbreaks, in our strict evaluation, this was not considered relevant.

$$Recall = \frac{\text{number of relevant items returned}}{\text{number of relevant articles}} \quad Precision = \frac{\text{number of relevant items returned}}{\text{number of items returned}} \tag{1}$$

Recall is based on the level of relevance of the PADI-web articles; i.e., an article is considered very relevant (++) if its publication date is before the date of the report or official notification (early identification of the outbreak) and is considered relevant (+) if its publication date is the same as or later than the date of the report or official notification. Two values of recall (R1 (++) and R2 (+ and ++)) and precision (P1 (++) and P2 (+ and ++)) were calculated. The harmonic mean of the precision and recall was also calculated (F-measure).

Note that the method used to compute precision is strict because relevant outbreaks correspond to the expected outbreaks for each study case. However, if PADI-web returns relevant outbreaks that are not the expected outbreaks related to our study cases, these could be considered relevant items (i.e., outbreaks found in the WAHIS database) in another measure of precision (called "real precision").

Note that the calculation of real and strict precision values as well as recall is based on manual analysis of the relevance of 244 outbreaks returned by our strategies.

Table 1. Precision, Recall, and F-measure for each strategy.

	Strategy A	**Strategy B**	**Strategy C**	**Strategy D**	**Strategy E**
R1	61.1%	31.3%	38.9%	**94.4%**	38.9%
Strict P1	17.7%	10.4%	**22.6%**	14.4%	18.4%
Strict F-measure (++)	27.5%	15.6%	**28.6%**	25%	25%
R2	59.7%	46%	41.8%	**94%**	44.8%
Strict P2	32.8%	28.2%	**48.3%**	28.6%	39%
Strict F-measure (+ and ++)	42.3%	35%	**44.8%**	43.9%	41.7%
Real Precision	69%	67%	**72.4%**	61.4%	66%
Real F-measure	64%	54.6%	53%	**74.3%**	53.4%

Table 1 shows better recall for Strategy D (and good results for Strategy A) and better precision for Strategy C. Strategy C highlights only events at the beginning of the articles. This explains why some relevant events were not extracted and why there were lower values of recall.

In terms of the F-measure, the best strategies are C and D, which are based on "strict" and "real" calculus methods, respectively. In this context, these strategies will be integrated into the MUST-AI import system.

4.3. Discussion

Many solutions exist for visualizing multisource data based on technical tools such as Elasticsearch/Kibana[7]. Therefore, several visualizers and dashboards have been developed over the past few years for epidemiological purposes, especially since the COVID-19 pandemic. They mostly focus on the number of confirmed cases represented on maps and incidence curves used to visualize spatiotemporal distributions and trends. More sophisticated platforms combine multisource information such as big data mining, remote sensing images, population data and other hazard data (disasters, chemical or nuclear expositions, etc.) [8, 9]. Although these solutions provide valuable help for

[7] https://www.elastic.co/

synthesizing large volumes of highly disparate data, the main challenge remains the integration of tools into routine workflows, sustainability and interoperability with existing information systems [4]. The animal health sector still needs substantial improvements in data-driven and web-based surveillance to bring prevention strategies into the digital age [5].

Some event-based surveillance tools, such as HealthMap, GPHIN or EIOS, exist, but they partially cover animal health issues or are not publicly available. Other tools exist to address specific animal health events, such as SARS-Ani. This tool was developed to characterize SARS-CoV-2 cases confirmed in nonhuman species reported in WAHIS and ProMED. It provides valuable insights into patterns of reporting and identifies data gaps between the official number of reports and known cases reported by other sources. However, the updating process requires manual data collection, resulting in a greater investment in human resources [12].

In MUST-AI, we focused on automatically collecting unstructured data from ProMED and PADI-web and extracting valuable information to make comparisons and complete official data. MUST-AI was designed to assist epidemic intelligence teams in producing early-warning messages and summaries. To our knowledge, there is currently no such automated tool publicly available for the detection of unusual health events in animals. The ”Highlight” option focuses on recent events and allows visualization of their locations and easy access to related information. Further developments will include filters on the dashboard to display specific categories (e.g., regions, groups of species, and sources) and a module to download data to perform more specific statistical analyses.

In this study, we identified two event extraction strategies that can be chosen according to the location of interest and the goal of the intelligence team. The most sensitive strategy (i.e., Strategy D) can be chosen when targeting countries with low investment in wildlife surveillance, when little is known about the disease situation, or even when targeting small countries (only a few articles are expected). In addition, the very relevant option (++) allows early warning signals to be identified when reporting issues are suspected. On the other hand, the most specific strategy (i.e., Strategy C) can be chosen to avoid ”noise” and target specific events. This approach could contribute to obtaining more information on the context and to complete sparse information content in official reports. It can also be used for large countries, where a large number of articles are expected. In further developments, users will be able to select the preferred strategy according to their goals.

Finally, research biases were encountered in the construction of the case studies, namely, that certain potentially relevant articles could not be studied because either the URL referring to the article on the web was obsolete or a subscription was required for access. In addition, PADI-web articles for which the publication date was too far from the date of the report or official notification were not considered during the study.

5. Conclusion and future work

MUST is an innovative surveillance tool that aims to combine the surveillance data from the official programs extracted from WAHIS and two web-based systems (i.e., ProMED-mail and PADI-web). In this paper, we described the first component of this tool, MUST-AI, which focuses on HPAI events detected in mammalian species. The events are displayed on a dashboard, which includes a map interface, an incidence graph, and the text of reports. This paper highlights the relevance of integrating events from Strategies C and D of PADI-web into MUST-AI.

In our future work, we plan to evaluate the quality of NER models (i.e., perform comparisons between generic and specific NER models learned from our labeled datasets) and the impact of the final results obtained.

This paper presents an evaluation of the integration of PADI-web data into MUST-AI, and we plan to conduct the same type of evaluation with ProMED data in future work.

Note that the same event can be represented in the MUST-AI map by all three data sources (WAHIS, ProMED and PADI-web) at the same time, causing the same event to appear several times (duplication). Therefore, these duplicates must be merged so that the event is represented only once in the MUST interface. This can be done by identifying criteria and fusion strategies to allow global visualization of HPAIM cases. Thus, an evaluation of fusion criteria based on the case study presented in this paper could constitute a continuation of this study.

Acknowledgments

This project received funding from the European Union's Horizon 2020 Research and Innovation Program under grant agreement No. 874850 and is cataloged as MOOD109. The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission. This work was also funded by the French General Directorate for Food (DGAL). We thank Paolo Tizzani from WOAH for his explanation of the WOAH reporting system and for the availability of the WAHIS data.

References

[1] Agüero, M., Monne, I., Sánchez, A., Zecchin, B., Fusaro, A., Ruano, M.J., del Valle Arrojo, M., Fernández-Antonio, R., Souto, A.M., Tordable, P., Cañás, J., Bonfante, F., Giussani, E., Terregino, C., Orejas, J.J., 2023. Highly pathogenic avian influenza a(h5n1) virus infection in farmed minks, spain, october 2022. Eurosurveillance 28. doi:https://doi.org/10.2807/1560-7917.ES.2023.28.3.2300001.

[2] Arsevska, E., Valentin, S., Rabatel, J., de Goër de Hervé, J., Falala, S., Lancelot, R., Roche, M., 2018. Web monitoring of emerging animal infectious diseases integrated in the French Animal Health Epidemic Intelligence System. PLOS ONE 13, e0199960. doi:10.1371/journal.pone.0199960.

[3] Carrion, M., Madoff, L.C., 2017. ProMED-mail: 22 years of digital surveillance of emerging infectious diseases. International Health 9, 177–183. doi:10.1093/inthealth/ihx014.

[4] Carroll, L.N., Au, A.P., Detwiler, L.T., chieh Fu, T., Painter, I.S., Abernethy, N.F., 2014. Visualization and analytics tools for infectious disease epidemiology: A systematic review. Journal of Biomedical Informatics 51, 287–298. doi:https://doi.org/10.1016/j.jbi.2014.04.006.

[5] Charlier, Johannes and Barkema, Herman W. and Becher, Paul and De Benedictis, Paola and Hansson, Ingrid and Hennig-Pauka, Isabel and La Ragione, Roberto and Larsen, Lars E. and Madoroba, Evelyn and Maes, Dominiek and Marin, Clara M. and Mutinelli, Franco and Nisbet, Alasdair J. and Podgorska, Katarzyna and Vercruysse, Jozef and Vitale, Fabrizio and Williams, Diana J. L. and Zadoks, Ruth N., 2022. Disease control tools to secure animal and public health in a densely populated world. LANCET PLANETARY HEALTH 6, E812–E824.

[6] Consortium, E., Flavia, O., Sascha, K., Carola, S.L., Christoph, S., Valerie, A., Alina, A., Sophia, B., Hannes, B., Caroline, B., Elena, B., Jiri, C., Nicolai, D., Friederike, G., Anja, G., Jörn, G., Moisés, G., Ignacio, G.B., Timm, H., Ferran, J., Oliver, K., Aleksija, N., Joaquin, N.H., Ilaria, P., Patricia, P.P., Jolianne, R., Katja, S., Tiziana, T., Kamila, P., Rachele, V., Gauthier, V., Natalie, W., Stefania, Z., Ezio, F., 2024. The role of mammals in avian influenza: a review. EFSA Supporting Publications 21, 8692E. doi:https://doi.org/10.2903/sp.efsa.2024.EN-8692, arXiv:https://efsa.onlinelibrary.wiley.com/doi/pdf/10.2903/sp.efsa.2024.EN-8692.

[7] Freifeld, C.C., Mandl, K.D., Reis, B.Y., Brownstein, J.S., 2008. HealthMap: Global Infectious Disease Monitoring through Automated Classification and Visualization of Internet Media Reports. Journal of the American Medical Informatics Association 15, 150–157. URL: https://doi.org/10.1197/jamia.M2544, doi:10.1197/jamia.M2544, arXiv:https://academic.oup.com/jamia/article-pdf/15/2/150/2086063/15-2-150.pdf.

[8] Goel, R., Valentin, S., Delaforge, A., Fadloun, S., Sallaberry, A., Roche, M., Poncelet, P., 2020. Epidnews: Extracting, exploring and annotating news for monitoring animal diseases. Journal of Computer Languages 56, 100936. doi:https://doi.org/10.1016/j.cola.2019.100936.

[9] Leung, C.K., Chen, Y., Hoi, C.S., Shang, S., Wen, Y., Cuzzocrea, A., 2020. Big data visualization and visual analytics of covid-19 data, in: 2020 24th International Conference Information Visualisation (IV), pp. 415–420. doi:10.1109/IV51561.2020.00073.

[10] Lin, S.Y., Beltran-Alcrudo, D., Awada, L., Hamilton-West, C., Schettini, A.L., Caceres, P., Tizzani, P., Allepuz, A., Casal, J., 2023. Analysing WAHIS Animal Health Immediate Notifications to Understand Global Reporting Trends and Measure Early Warning Capacities (2005–2021). Transboundary and Emerging Diseases ID 6666672, 10 pages. doi:10.1155/2023/6666672.

[11] Madoff, L.C., Li, A., 2014. Web-based surveillance systems for human, animal, and plant diseases. Microbiology Spectrum 2, 10.1128/microbiolspec.oh–0015–2012. doi:10.1128/microbiolspec.oh-0015-2012, arXiv:https://journals.asm.org/doi/pdf/10.1128/microbiolspec.oh-0015-2012.

[12] Nerpel, A., Käsbohrer, A., Walzer, C., Desvars-Larrive, A., 2023. Data on sars-cov-2 events in animals: Mind the gap! One Health 17, 100653. doi:https://doi.org/10.1016/j.onehlt.2023.100653.

[13] Paquet, C., Coulombier, D., Kaiser, R., Ciotti, M., 2006. Epidemic intelligence: a new framework for strengthening disease surveillance in Europe. Eurosurveillance 11, 5–6. URL: https://www.eurosurveillance.org/content/10.2807/esm.11.12.00665-en, doi:10.2807/esm.11.12.00665-en.

[14] Rabalski, L., Milewska, A., Pohlmann, A., Gackowska, K., Lepionka, T., Szczepaniak, K., Swiatalska, A., Sieminska, I., Arent, Z., Beer, M., Koopmans, M., Grzybek, M., Pyrc, K., 2023. Emergence and potential transmission route of avian influenza a (h5n1) virus in domestic cats in poland, june 2023. Eurosurveillance 28. doi:https://doi.org/10.2807/1560-7917.ES.2023.28.31.2300390.

[15] Valentin, S., Arsevska, E., Rabatel, J., Falala, S., Mercier, A., Lancelot, R., Roche, M., 2021. Padi-web 3.0: A new framework for extracting and disseminating fine-grained information from the news for animal disease surveillance. One Health 13, 100357. doi:https://doi.org/10.1016/j.onehlt.2021.100357.

[16] Vreman, S., Kik, M., Germeraad, E., Heutink, R., Harders, F., Spierenburg, M., Engelsma, M., Rijks, J., van den Brand, J., Beerens, N., 2023. Zoonotic mutation of highly pathogenic avian influenza h5n1 virus identified in the brain of multiple wild carnivore species. Pathogens 12. URL: https://www.mdpi.com/2076-0817/12/2/168, doi:10.3390/pathogens12020168.