We thank the anonymous referee 2 for their valuable comments on our manuscript, 'Leveraging Social Media for Disaster Management: A Critical Review of Data Collection Strategies and Actionable Insights' (egusphere-2024-1536, submitted to NHESS). We have carefully considered all the comments and will make revisions to address your suggestions. Below, we provide detailed responses to each of Referee 2's comments (R2).

**R2 Comment 1**: "*The manuscript presents a critical review of the use of social media data in the context of disaster management, reviewing a total of 250 articles from a range of disciplines. The paper is well-written and offers a valuable contribution to the growing literature on social media in disaster contexts. The paper would benefit from a clearer focus on its core research questions and a more concise presentation of the findings. The current version includes a significant amount of descriptive content, which could be streamlined to better highlight the most important insights*."

**R2-1 Reply**: Thank you for the summary and feedback. We truly appreciate your positive assessment of our manuscript and your recognition of its contribution to the literature on social media use in disaster management.
We also acknowledge your suggestion to sharpen the focus on the core research questions and streamline the descriptive content. Accordingly, we will revise the manuscript to present the findings more concisely and reduce sections with excessive descriptive content.

---

**R2 Comment 2**: "*Length and focus of the study: The manuscript is quite long (44 pages, including 19 figures and 4 tables). While comprehensive, much of the content is descriptive and does not directly address the core research questions. I recommend moving some of the less critical descriptive sections (particularly in Sections 3.3, 4.2, and 4.4) to the appendix. For me, the most insightful sections were 4.3 and the discussion in Section 5. A focus on these could improve the impact of the paper.*"

**R2-2 Reply**: Thank you for this constructive feedback regarding the length and focus of the manuscript. We appreciate your suggestion to streamline descriptive sections and highlight the most impactful insights.
Regarding Section 3.3, which describes the synthesis of research findings and details each column of the Social Media Literature Database, we agree that the full content may be too detailed for the main text. However, we believe it is important to retain a summarized version in the main manuscript to help readers understand the structure and reasoning behind the database design. We will revise this section to briefly summarize the main categories and subcategories in the paper, and will move the detailed descriptions of each subcategory to the appendix.
For Section 4.2, which presents early works, we have opted to relocate most of the content to the appendix, as suggested. However, a brief overview will be retained in the main manuscript (added to section 2 - Background) to preserve continuity and context for readers unfamiliar with the background.
With regard to Section 4.4, we believe this section is directly relevant to the paper's goals, as it provides insights into how disaster management research is distributed across themes and publication sources. It

also supports the structure of our classification system. Therefore, we have chosen to retain this section in the main manuscript but will tighten the narrative to improve conciseness and clarity.

---

**R2 Comment 3**: "*Clarification of research questions and key insights: The paper could more clearly articulate its central research questions and ensure that the findings directly address them. In particular, I found the second research question on the "actionable information" derived from social media more relevant and of interest to a broader audience. In contrast, the first question on "exclusion criteria in relevance filtering" is more technical and may be of limited interest to non-specialist readers.*"

**R2-3 Reply**: Thank you for highlighting the importance of clearly articulating the research questions and ensuring that the findings directly address them. We agree with your observation that Research Question 2, which focuses on "actionable information" derived from social media, has strong relevance to a broader interdisciplinary audience and have emphasized this further in our revisions.

Regarding Research Question 1 on "exclusion criteria in relevance filtering," we acknowledge that it is more technical in nature. However, we would like to emphasize that it addresses a critical aspect of working with social media data. Practically all researchers working with such data, regardless of domain, engage with some level of technical processing, such as querying APIs, parsing noisy content, and filtering for relevance. In such a vast and unstructured space, the ability to identify and exclude irrelevant data is not only technical but foundational to meaningful analysis. Exclusion using keyword-based filtering, NLP techniques, or ML methods is a common and necessary practice across many applications.

To make this more accessible, we will revise Section 3.1 to better explain the relevance and motivation behind Research Question 1 in a way that is understandable to a wider audience. We will also rewrite Section 4.8, where the results of this question are presented, with clearer framing and explanation, so its importance is evident even to those outside the technical field.

---

**R2 Comment 4: "***Practical examples in the introduction: The introduction would benefit from one or two concrete examples illustrating the kind of information that can be derived from social media in disaster contexts and how these tools have been used by researchers. This would help orient readers who are less familiar with the data sources or their potential applications.*"

**R2-4 Reply**: We agree that including examples of how social media data has been used in disaster contexts would greatly enhance the clarity and accessibility of the introduction, particularly for readers who may be less familiar with the data sources or their applications.

In response, we will revise the introduction to include brief practical examples. Specifically, we will mention how social media was used for real-time situational awareness during events such as Hurricane Sandy, where tweets were analyzed to detect flooded streets and power outages. We will also highlight how user-generated content from platforms like Twitter and Facebook has been used to assess infrastructure damage and coordinate relief efforts.

---

**R2 Comment 5: "***Keyword selection: The search strategy seems to have focused on Twitter, with no mention of other major platforms such as Facebook, Instagram, or Weibo. This is particularly surprising*

*given that some of these platforms have significantly higher user bases and broader geographic reach. They have also been extensively used in disaster contexts (e.g., Facebook). The rationale for this focus should be clearly explained, and the implications of this potential bias should be acknowledged more explicitly in both the methodology and the interpretation of findings (e.g., when the authors find that the great majority of studies in their database uses Twitter as noted in line 338 and elsewhere).*"

**R2-5 Reply:** Thank you for highlighting this important point. A similar comment was raised by Referee 1, and we have provided a detailed response and corresponding revisions to address this concern.
In summary, we have clarified in Section 3.2 (Paper Searching Criteria) that the search strategy focused on Twitter due to its relatively open API access, which has historically made it more accessible for data collection in disaster management research. We now explicitly acknowledge the implications of this platform preference in our methodology and in the interpretation of our findings (Sections 5.7). Additionally, we conducted a comparative keyword search using other platform names and included a summary table to contextualize the potential bias.

---

**R2 Comment 6**: "*Inclusion of studies not using social media: Section 3.3.4. describes the data used in studies distinguishing whether article have utilized social media data or not. It is unclear why studies not using social media were included in the database, given the paper's stated focus on social media. Of course, the authors could use these other articles as reference points to show the advantages or disadvantages of the use of social media, but this does not really happen in the analysis. Instead, when describing their database, the authors also refer to those articles not using social media data, which may confound the analysis.*"

**R2-6 Reply**: We acknowledge the confusion this may have caused and agree that clarification is necessary.
We have included a small number of articles in the database that do not directly utilize social media data for two main reasons: (1) some of these are review or survey papers that critically discuss the role and application of social media in disaster management, even though they do not perform direct data collection or analysis; (2) others present methods for social media data collection or relevance filtering, but apply them only conceptually or use proxy data (e.g., user-generated content from official platforms), rather than directly collected social media data.
To address this, we will revise Section 3.3.4 and other relevant parts of the manuscript to clearly state the rationale behind including such studies and distinguish their role within the database.

---

**R2 Comment 7**: "*Stronger emphasis on key messages: While the paper shows great technical detail, the manuscript would benefit from a more focused discussion of key takeaways. Specifically, what are the major advantages, disadvantages, and challenges of using social media data in disaster contexts? These insights could be more prominently featured in the abstract, conclusion, and discussion sections (see also my comment 2 on the research question above).*"
**R2-7 Reply:** We agree that clearly emphasizing the key takeaways, advantages, and disadvantages would enhance the impact of the manuscript.

Advantages of using social media in disaster contexts include the availability of real-time, user-generated content from affected individuals and eyewitnesses, often enriched with metadata (e.g., time and location), which can significantly aid rapid response and decision-making.
Challenges include the reliability and credibility of content (e.g., misinformation, unverifiable posts), and the informal nature of language used, such as abbreviations, slang, or multilingual content, which complicates automated analysis and necessitates the use of NLP and ML methods.
Actionable Information: Our review identifies that social media data can provide geolocation detection, community collaboration signals, and disaster trends or hotspot identification. These capabilities directly support operational decision-making during and after a disaster, highlighting the potential of social media as a critical data source.
In relation to Research Question 1, we reinforce that keyword-based exclusion criteria form a technically sound starting point, especially considering that most social media APIs are built on keyword filters. However, for more advanced and adaptive filtering, we recommend NLP-based approaches and ML models, which can help reduce noise and improve the relevance of collected data.
We will revise the abstract, conclusion, and discussion sections to more directly highlight these insights of our review.

---

**R2 Comment 8**: "*Line 35: The sentence 'Existing literature reviews on social media data (SMD) platform evaluations, data collection tools, and analysis methods over time' is missing a verb.*"
**R2-8 Reply:** 'There are several existing literature reviews on social media data (SMD) platform evaluations, data collection tools, and analysis methods over time.' This revision corrects the grammatical issue and clearly states the existence of prior reviews.

---

**R2 Comment 9**: "*Line 176: Typo in "The may however be other keywords..."*
**R2-9 Reply:** 'There may, however, be other keywords that we did not use which would have identified further relevant literature.' This revision corrects the typo.

---

**R2 Comment 10:** "*Line 180ff: The authors note that relevant studies may have been missed. Would it not have been advisable to revise the keyword strategy to more comprehensively capture the literature?*"
**R2-10 Reply:** We agree that refining the keyword strategy could have led to a more comprehensive capture of the literature. However, as part of our critical review methodology, we proceeded with analyzing the 250 articles identified in the first phase (Phase I) of our review, which used the Boolean search strings initially developed based on our topics of interest.
In Lines 180–185 under Section 3.2, we explicitly acknowledge this as a potential limitation and highlight it as a methodological bias. We intended to maintain transparency about the scope of our keyword strategy and its influence on the literature selection. In future studies, we plan to incorporate an iterative keyword refinement process to broaden coverage. This point will be clarified in the revised manuscript.

---

**R2 Comment 11:** "*Line 288: Missing word: 'Such data is considered credible when compared to the data extracted from social media platforms as it may include false information.*"

**R2-11 Reply:** The updated version now reads: “Such data is considered more credible compared to data extracted from social media platforms, which may include false or misleading information.”

---

**R2 Comment 12:** “*Section titles: Some subsection titles could be made more meaningful and specific. For example, “4.2. Early Works” could clarify the time period; “4.2.2. Previous Works on Social Media Analytics” could clarify what is meant with “previous”, etc.*”

**R2-12 Reply:** The section titled “4.2 Early Works” refers to studies published between 2010 and 2023; we will update the title to reflect this time frame explicitly. Similarly, “4.2.2 Previous Works on Social Media Analytics” also refers to early studies within the same period (2010–2023). We will revise the heading to avoid ambiguity. Additionally, we will review other section titles across the manuscript to enhance clarity.

Thank you for your suggestions, and we hope the above-mentioned revisions will meet your expectations.

********