We thank the anonymous referee 1 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 1 comments (R1).

R1 comment 1: "*The authors present a literature review and analysis with an associated dataset of 250 articles, that they systematically retrieved to categorize concerning study area, event, data details, and methods. The authors emphasize that the evaluated studies consistently show that social media facilitates community interactions in crisis and that the main remaining concern is assuring accuracy by addressing the unreliability of the data. They specifically focus on actionable insights from the reviewed papers and present the results from the categorized literature with a number of well-compiled images. The insights they draw from the literature are timely and represent a novel contribution and a future reference for the field of social media usage and analysis during natural hazards.*"

R1-1 reply: We thank R1 for their summary and comments, recognising the timeliness and contribution our critical review makes.

R1 comment 2: "*Nevertheless, I have identified one point that requires improvement/clarification in the interpretation of their findings and a few minor suggestions for the manuscript. Figure 3 shows the search term combinations, which were used to generate the database. each of the search strings contains a word related to the platform Twitter, but no other commonly used platforms (e.g. Weibo, Facebook, Instagram TikTok, and more) are included in the search. Since the Twitter API did provide researchers with free access, this search terminology is valid and it is very transparently reported. However, in L338-L339 and L685-690 the authors report a platform bias toward Twitter. This might be a trend in research, but I think the dataset is not suited to underline this finding. If the authors report this bias on the basis of a dataset, that was filtered articles with the word "Twitter", naturally the authors will retrieve more articles, that base their analysis on data from the platform Twitter and they should not report this as a representative result. I would assume that if the search term Twitter was replaced with the search term "Weibo" we might have a much higher percentage of articles using this platform (although probably Twitter would still be the number one). So here I would expect this limitation to be mentioned alongside the results and probably mention platform-related results with higher uncertainty*"

R1-2 reply: We thank R1 for their comment regarding the Boolean search strings, noting that only the keyword Twitter is used and that no other commonly used platforms are included in the search. As you have mentioned in the comment, Twitter (currently X) data was more widely available to researchers than other platforms. The most popular social media platforms, such as Weibo, Facebook, Instagram, TikTok, Snapchat, YouTube, Telegram, and WhatsApp, do not make their crowdsourced data available freely to the public and hence they are not used much in social media and disaster management research.

To expand on this point, and we agree this was not clarified in our original manuscript, we will revise the manuscript to include the following clarifications:

1. Acknowledgment of only using Twitter in Search Terms

    Under section 3.2 (Paper Searching Criteria). We will mention that our search strings only used Twitter as a social media platform and that excluding other commonly used platforms such as Weibo, Facebook, Instagram, and TikTok may have omitted relevant research from these platforms. However, we will note that around 5% of the articles we found using the search strings that have just Twitter and included in our database mention in their articles the use of social media data from other platforms (discussed in sections 4.2.2 and 5.7).

2. Discussion of methodological social media biases used by researchers.

    Table 1 below shows the usage of social media platforms (most common platforms, excluding personal chat platforms such as Whatsapp or Telegram) used in disaster management research in addition to Twitter. We reran our original search strings with the term 'Twitter' and then replaced the term "Twitter" in each search string (Q1, Q2, Q3, Q4, Q5) with six different social media platforms. The column "Results from Google Scholar" represents the number of articles filtered by Google Scholar using the search query. The column "Articles related to the social media platform" represents the number of articles that specifically use the respective platform, as determined by reading the abstract and title of each article.

The search query references conducted in Google Scholar (English, to present) are as follows:
- *Q1: allintitle: ("social media" OR "twitter") AND ("Disaster Response" OR "Disaster Mitigation" OR "Disaster Recovery" OR "Disaster Preparedness" OR "Disaster Monitoring")*
- *Q2: allintitle: "disaster management" AND ("social media" OR "twitter" OR "news" OR "crowdsourcing")*
- *Q3: allintitle: "data collection" AND ("disaster" OR "hazard" OR "flood" OR "landslide") AND ("social media" OR "twitter" OR "tweet")*
- *Q4: allintitle: ("social media" OR "twitter") AND ("disaster" OR "hazard") AND ("data" OR "filtering" OR "exclusion")*
- *Q5: allintitle: ("social media" OR "twitter") AND ("emergency response" OR "disaster relief")*

Table 1. *Comparison of X (Twitter) Google Search results using search strings Q1 to Q5 with replacing the word "Twitter" with other social media platforms (Facebook, Weibo, Instagram, TikTok, Reddit, Quora). Section A of the table are all results from Google Scholar, and Section B, the results after examining abstracts and titles for relevance. Note that the same peer-review article might appear under different rows.*

Social Media Platform	No. of articles (a = original analyses Jan 2010 to Sep 2023; b = new analyses Jan 2010 to July 2024).									
	A. Results from Google Scholar using search strings Q1 to Q5 which includes the social media platform					B. Articles related to the social the given media platform (after examining abstract and title)				
	Q1	Q2	Q3	Q4	Q5	Q1	Q2	Q3	Q4	Q5
X (Twitter) (a)	107	125	4	81	48	82	112	2	23	31
X (Twitter) (b)	123	145	5	117	79	85	117	5	38	34
Facebook (b)	122	135	3	88	68	8	3	0	2	4
Weibo (b)	115	131	3	88	65	5	1	0	2	2

Instagram (b)	116	134	3	86	65	1	2	0	0	0
TikTok (b)	115	132	3	86	65	0	0	0	0	0
Reddit (b)	115	132	3	86	65	0	1	0	0	1
Quora (b)	115	132	3	86	65	0	0	0	0	0

In section 5.7 (Methodological Biases), we will therefore discuss these results and add a discussion point that elaborates on the fact that Twitter has historically provided more accessible data through its API compared to other platforms.

3. Highlighting Uncertainty in Platform-Related Results
   We will revise the interpretation of our findings in L338-L339 and L685-690 to include a statement about the uncertainty associated with platform-related results. We will clarify that our search term selection influences the observed platform bias towards Twitter and may not fully represent the overall trend in disaster management research.

   In L338-L339 and L685-690, we have noted that Twitter is the preferred social media platform for disaster management research primarily due to the accessibility of its data compared to other platforms. We recognize that reporting this as a bias is not entirely appropriate since our search strings specifically included the term "Twitter," which influenced the dataset composition.

   We still believe it is important to discuss the platform preference bias from a broader perspective that includes the overall literature, not just our dataset. While our search strings were tailored to include "Twitter," this preference reflects a trend within the research community where Twitter data is more readily available and thus more frequently used.

   We will revise the manuscript to clarify this point (section 5.7), ensuring that the platform-related results are presented with higher uncertainty and explicitly acknowledging the limitation introduced by our search criteria.

R1 Comment 3: "*Miner improvements I would suggest are -*
L52: "section 1" → "Section 1";
L97 "delves into" L310 "delving into" L500 "delve into" → this wording is overused by LLMs recently and therefore I would change;
L176: "The may…" → "There may…";
L198: I think the commas should be outside the quotation marks;
L283: Citation is all caps, please correct, also in following instances;"

R1-3 Reply: The changes have been made in the manuscript.

R1 Comment 4: "*L370/375: double association of the acronym AI for actionable insights and Artificial Intelligence is confusing, please use a different one*"

R1-4 Reply: A different acronym will be given for the term 'Actionable Information' in the manuscript. A few possible acronyms are 'ActInfo' and 'ActIn' which will be finalised post internal discussions with the authors.

R1 Comment 5: "*Figure 1: Box "Research Question Identification" → Q1 is unclear to me*"

R1-5 Reply: Figure 1 represents a flow diagram of our review process and the first box "Research Question Identification" explains the main 2 research questions we have focused on in our work. RQ1 states "Does exclusion criteria assist in relevance filtering?" and is elaborated in section 3.1. Exclusion criteria are NLP-based programmatic conditions that can be used to eliminate noise in data. This method is widely used and we seek to identify how the methodologies in the literature have used this method to filter data. We are trying to find out if using exclusionary criteria in data collection methodologies contributes to noise removal, which ultimately aids in rapid decision-making. We will modify the figure caption by mentioning which section to refer to (3.1) to read the explanation of the research questions.

R1 Comment 6: "*Figure 14: Why is some text in black and some in white? This makes the Figure more confusing.*"

R1-6 Reply: The colours were chosen for the text for better readability, and can be modified to reduce confusion. Either the boxes written in grey-black can be modified to white colour. If readability is affected, every box caption will be modified to grey-black.

R1 Comment 7: "*Figure 15 - The representation of quantities in lakh is an Indian numbering system, that is not commonly used in many other countries and therefore not very suited for an international audience.*"

R1-7 Reply: The quantities in the figure will be modified to millions using the conversion (1 lakh = 100,000).

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

***