Centre for Research on the Epidemiology of Disasters CRED

UCLouvain

USAID FROM THE AMERICAN PEOPLE

EM-DAT
Scientific Technical Advisory Group Meeting (STAG)
of
the International Emergency Events Database (EM-DAT)

Brussels, March 20-21, 2023.

Meeting Notes

Acronyms – a selection

- **CIRED :** Centre international de recherche sur l'environnement et le développement
- **CRED:** Center for Research on Epidemiology and Disasters
- **EM-DAT:** The International Disaster Database
- **IFRC:** International Federation of the Red Cross
- **IRDR:** Integrated Research on Disaster Risk
- **IRSS:** Institute for Research on Health and Society (UCLouvain)
- **NOAA:** National Oceanic and Atmospheric Administration
- **PDNA:** Post Disaster National Assessment (per sector, per country, per event)
- **STAG:** Scientific and Technical Advisory Group
- **BHA (USAID):** Bureau of Humanitarian Affairs
- **UNDRR:** United Nations Office for Disaster Risk Reduction
- **WMO:** World Meteorological Organization

List of Participants

Ece AKSOY
Food and Agriculture Organization of the United Nations (FAO), Italy

Lucia BEVERE
Swiss Re, Switzerland

Rhonda STEWART
United States Agency for International Development (USAID), US

James DOURIS
World Meteorological Organization (WMO), Switzerland

Justin GINNETTI
International Federation of Red Cross and Red Crescent Societies (IFRC), Switzerland

Aglaé JÉZÉQUEL
Laboratoire de Météorologie Dynamique (LMD), France

Nathaniel KAMMER
United States Agency for International Development (USAID), US

Matthieu KERVYN
Vrije Universiteit Brussel (VUB), Belgium

Albert KETTNER
Dartmouth Flood Observatory (DFO), US

Andrzej KIJKO
University of Pretoria, Natural Hazard Centre, South Africa

Seppe LAMPE
Vrije Universiteit Brussel (VUB), Belgium

Petra LÖW
Munich Re, Germany

Virginia MURRAY
UK Health Security Agency, United Kingdom

Ilan NOY
University of Victoria, Wellington, New Zealand

Sylvain PONSERRE
Internal Displacement Monitoring Centre *(*IDMC), Switzerland

Bishal RAJ GURUNG
National Society for Earthquake Technology, Nepal

Adam ROWLAND FYSH
United Nations Office for Disaster Risk Reduction (UNDRR), Switzerland

Tahmina SHIRIN
Institute of Epidemiology, Disease Control and Research, Bangladesh

Adam SMITH
National Oceanic and Atmospheric Administration *(*NOAA), US

Sezin TOKAR
United States Agency for International Development (USAID), US

Iria TOUZON CALLE
United Nations Office for Disaster Risk Reduction (UNDRR), Germany

Sandy TUBEUF
Institute Health and Society (IRSS), UCLouvain, Belgium

Marnik VANCLOOSTER
UCLouvain, Belgium

Joris VAN LOENHOUT
Sciensano, Belgium

Dewald VAN NIEKERK
North-West University, African Centre for Disaster Studies, South Africa

Aleeza WILKINS
United States Geological Survey (USGS), US

Paola YELA BELLO
International Federation of Red Cross and Red Crescent Societies (IFRC), Switzerland

Zehra ZAIDI
Food and Agriculture Organization of the United Nations (FAO), Italy

Regina BELOW; Damien DELFORGE; Géraldine DELESTIENNE; Rebecca JONES; Cinzia LANFREDI SOFIA; Debby PARAMITASARI; Niko SPEYBROECK; Margo TONNELIER; Valentin WATHELET
Centre for Research on the Epidemiology of Disasters (CRED) team, Institute Health and Society (IRSS), UCLouvain, Belgium

Agenda

Day 1 – 20th March 2023 a.m.

08:20 – 08:45	Arrival + Registration
08:45 – 09:30	*Welcome:* Niko Speybroeck (UCLouvain) – Rhonda Stewart (BHA/USAID) Meeting objectives Icebreaker session
09:30 – 10:00	*Presentation:* Niko Speybroeck (UCLouvain). *Ongoing work related to EM-DAT*
10:00 – 10:20	Tea & coffee break
10:20 – 10:50	*Plenary session*: Classification of disasters, definitions and terminology Moderator: James Douris (WMO) *Presentation:* Damien Delforge (UCLouvain). *Classification of Hazards & Disasters: The EM-DAT Perspective*
10:50 – 11:45	*Breakout session 1: Classification of disasters, definitions and terminology* Group 1: UN and international organizations Lead: Sezin Tokar (USAID) – Rapporteur: Valentin Wathelet Group 2: Research and academia Lead: Matthieu Kervyn (VUB) – Rapporteur: Rebecca Jones Group 3: Global data Lead: Lucia Bevere (Swiss Re) – Rapporteur: Cinzia Lanfredi
11:45 – 12:30	*Re-group + review conclusions breakout session 1 groups* Moderator: James Douris (WMO)
12:30 – 13:30	Lunch

Day 1 – 20th March 2023 p.m.

	Expert Presentations Moderator: Joris Van Loenhout (Sciensano)
13:30 – 13:45	Justin Ginnetti (IFRC). *The Global Crisis Data Bank, a multi-stakeholder initiative.*
13:45 – 14:05	Adam Smith (NOAA). *U.S. Billion-dollar Weather and Climate Disasters: Data Tools, Multi-Hazard and Socioeconomic Risk Mapping.*
14:05 – 14:20	Adam Rowland Fysh (UNDRR). *Risk Information Exchange and Classifying Hazards.*
14:20 – 14:35	Iria Touzon Calle (UNDRR). *New disaster damages and losses tracking system for countries: update and way forward.*
14:35 – 14:50	Aglaé Jézéquel (LMD). *Extreme heatwaves in Europe 1950-2020.*
14:50 – 15:05	Sylvain Ponserre (IDMC*). Disaster Displacement, Informing action on internal displacement with data, research and evidence.*
15:05 – 15:30	Tea & coffee break
	Expert Presentations Moderator: Joris Van Loenhout (Sciensano)
15:30 – 15:45	Dewald Van Niekerk (African Centre for Disaster Studies). *Use of EM-DAT to evaluate the Sendai Framework for Disaster Risk Reduction in Africa*.
15:45– 16:00	Virginia Murray (UK Health Security Agency). *UNDRR-ISC Hazard Information Profiles and their use in documenting hazards.*
16:00 – 16:15	Paola Yela Bello (IFRC). *Alerting the general public to hazards: overview of the IFRC CAP editor.*
16:15 – 16:30	Zehra Zaidi (FAO). *Representing Agricultural Losses in Disaster Data*.
16:30 – 17:00	*Wrap-up of Day 1* Niko Speybroeck (UCLouvain) – Rhonda Stewart (BHA/USAID)
18:30	Dinner

Day 2 – 21st March 2023 a.m.

09:00 – 09:15	*Arrival*
09:15 – 09:30	*Recap of Day 1* – Niko Speybroeck (UCLouvain)
09:30 – 11:00	*Plenary session*: Tackling data quality issues Moderator: Sandy Tubeuf (UCLouvain) *Presentations:* Rebecca Jones (UCLouvain). *Missingness in Global Disaster Data* Damien Delforge (UCLouvain). *Biases and Missingness in EM-DAT Disaster Data: Recent Technological Opportunities* Discussion
11:00 – 11:15	Tea & coffee break
11:15 – 12:00	*Breakout session 2: Tackling data quality issues* Group 1: Handling missing data Lead: Adam Smith (NOAA) – Rapporteur: Cinzia Lanfredi Group 2: Automating data collection Lead: Sylvain Ponserre IDMC – Rapporteur: Valentin Wathelet Group 3: Improving data quality, including epidemics and heatwaves Lead : Aglaé Jézéquel (LMD) – Rapporteur : Margo Tonnelier
12:00 – 12:45	*Re-group + review conclusions breakout session 2 groups* Moderator: Petra Löw (Münich Re)
12:45 – 13:45	Lunch

Day 2 – 21st March 2023 p.m.

13:45 – 14:00	*Plenary session*: Looking to the future: priorities and recommendations for EM-DAT: introduction Moderator: Niko Speybroeck (UCLouvain) and Rhonda Stewart (USAID)
14:00 – 15:00	*Breakout session 3: Priorities and recommendations for EM-DAT* Group 1: UN and international organizations Lead: Adam Rowland Fysh (UNDRR) – Rapporteur: Rebecca Jones Group 2: Research and academia Lead: Ilan Noy (University Victoria, NZ) – Rapporteur: Margo Tonnelier Group 3: Global data Lead: Dewald Van Niekerk (African Centre for Disaster Studies) – Rapporteur: Debby Paramitasari
15:00 – 15:30	Tea & coffee break
15:30 – 16:15	*Re-group + review conclusions breakout session 3 groups* Moderator: Petra Löw (Münich Re)
16:15 – 16:45	*Wrap-up of Day 2* Niko Speybroeck (UCLouvain) – Rhonda Stewart (BHA/USAID)
16:45	Meeting closure and goodbye

Background, Meetings Objectives, and Organization

The inaugural Scientific and Technical Advisory Group (STAG) meeting, in conjunction with the activities of the Emergency Events Database (EM-DAT) project, took place on March 20 and 21, 2023, in Brussels. The Thon Louise Hotel in Brussels served as the venue for this two-day event.

The meeting was organized by the Centre for Research on the Epidemiology of Disasters (CRED), located at UCLouvain, and supported by the Bureau for Humanitarian Assistance of the United States Agency for International Development (BHA/USAID).

STAG's principal objective was to gather a select group of scientific experts tasked with providing advisory input on the quality, performance, and future potential enhancements of the EM-DAT data. The broader aim of the meeting was to provide strategic direction and identify potential areas of research for the EM-DAT project.

The STAG meeting convened specialists from diverse backgrounds, including representatives from academic institutions, research centers, disaster data management entities, and both local and international stakeholders.

Three main areas of discussion were the following:

- Disaster type definition and classification
- Automation processes for disaster data collection
- Methods to address data gaps

A significant portion of the meeting time was allocated to presentations and dialogues concerning the current progress of EM-DAT and the three main aforementioned discussion points. These subjects were further scrutinized in breakout group sessions that were divided either by organization type - (1) United Nations and international organizations; (2) Research and academic institutions; (3) Global data - or by discussion topics - (1) Addressing missing data; (2) Automating data collection; (3) Improving data quality, which includes data on epidemics and heatwaves.

Each session started with a comprehensive presentation, paving the way for subsequent discussions. Attendees were asked to summarize the key points from their discussions to facilitate the collection of suggestions for the future improvement of EM-DAT.

This document provides a summary of the meeting's proceedings. For more details, please follow this link: https://www.emdat.be/stag.

Day 1 – 20th March 2023

Opening

The first day of the meeting started with a welcoming address by Rhonda Stewart of BHA/USAID and Niko Speybroeck of CRED/IRSS/UCLouvain, hereby inaugurating the first Scientific and Technical Advisory Group (STAG) event.

A comprehensive overview of the meeting's objectives was provided, with a particular emphasis on how STAG meetings diverge from the previous Technical Advisory Group (TAG) meetings. The uniqueness of the STAG meeting lies in its assembly of scientific experts, thus encouraging scientifically-driven enhancements.

The meeting was marked by the successful assembly of experts from an array of international organizations and academic institutions. The objective of the STAG initiative is to consistently harness the expertise of STAG members as needed and the insights gained from this first STAG meeting will serve to guide the planning of subsequent meetings.

The opening session of the STAG meeting concluded with a roundtable session, enabling participants to introduce themselves and establish a collaborative environment.

In the closing part of this initial session, N. Speybroeck presented an overview of the current EM-DAT-related activities to ensure that the participants had a thorough understanding of the state of affairs before proceeding to subsequent presentations and discussions.

Plenary session: Classification of Disasters, Definitions, and Terminology

Presentation

Presenter: Damien Delforge (CRED, UCLouvain)

The Emergency Events Database (EM-DAT) has been an active, historical disaster database since 1988, playing a pivotal role in the research and study of natural hazards and disasters. It strictly includes disasters that meet a set of defined criteria and excludes conflicts or socio-economic-political events, even if these can cause significant losses to life and property.

Over the years, EM-DAT has undergone several enhancements, primarily due to collaborative network meetings like the past Technical Advisory Group (TAG) meetings. These gatherings have expanded the knowledge base and widened the community practicing disaster data collection and analysis.

A significant upgrade to EM-DAT's classification system was derived from discussions during the Integrated Research on Disaster Risk (IRDR) initiative in 2014. This initiative, led by a diverse group of stakeholders from the research community, government, United Nations agencies, humanitarian organizations, and private reinsurance companies, resulted in the IRDR Peril Classification and Hazard Glossary[1]. The glossary classified natural hazards into six main groups, providing a foundational structure for EM-DAT's current classification system. EM-DAT still relies on its historical definitions for its classification of natural and technological hazards.

In EM-DAT, disasters are classified based on their triggering hazard event – the extreme physical phenomenon (e.g., flood, drought, or explosion) causing the impact (fatalities, affected people, economic losses). If a disaster fulfills EM-DAT's inclusion criteria, its class is determined using EM-DAT's resources and a classification tree based on the IRDR reference. Additionally, EM-DAT uses an

[1] https://www.irdrinternational.org/knowledge_pool/publications/173

'associated disasters' system. This secondary classification system tags additional hazards that occurred concurrently or due to the main disaster. This system offers flexibility and does not strictly adhere to EM-DAT's primary classification tree.

Experience has taught EM-DAT that a robust classification system needs to be comprehensive, clear, consistent, suitable, stable, and adaptable, considering the varied perspectives of its user base. This balance can be challenging to achieve but is crucial for the database's continual evolution, particularly as the societal understanding of disaster risks grows. In practice, finding the right equilibrium will be a matter of compromise. The STAG may play an instrumental role in this evolution.

With its limited resources, knowledge, and perspective, the Centre for Research on the Epidemiology of Disasters (CRED) does not aim to overcome all challenges independently. It is worth noting that the STAG meeting underscored the classification system's inherent ambiguity, arising from varied user needs, diverse expertise, and the contrast between scientific consensus on disaster taxonomy and terminology used in mainstream media.

Breakout Session 1

Following the introductory presentation, participants engaged in breakout sessions that centered on the topics of disaster categorization, definitions, and terminology. The insights derived from these interactive discussions can be organized into four core themes: i) disaster classification, ii) data accuracy and user interaction, iii) re-classification and terminology refinements, and iv) challenges and recommendations. Table 1 summarizes the main insights.

Table 1: Breakout Session 1 'Classification of Disasters, Definitions, and Terminology', Main Insights

Disaster Classification
• Utilizing the UNDRR classification increases accuracy across systems. Maintaining a 'flat classification' at the EM-DAT level remains valid. Note: the hazard list is flat when not hierarchical, recognizing that a hierarchical classification may not adequately capture the complex interplay between different hazards. However, to aid usability, hazards are presented in a grouped structure with hazard types and hazard clusters. The clustering of hazards is not to be prescriptive as to the relationships of one hazard to another hazard or to many hazards. • An EM-DAT disaster classification should try to find a balance between the criteria of flexibility, transparency, and reproducibility. • Expansion of category levels or divisions of sub-groups and sub-types of hazard categories was discussed, with storms and associated hazard clusters as an example.
Data Accuracy and User Interaction
• A request was made for more clarity on attribution to improve EM-DAT, including details on metadata source labels, definitions, domain value, start and end dates of values, and attribute value accuracy. • The necessity for understanding the user base beyond academic needs was emphasized, particularly focusing on local authorities and other institutions. • A glossary of terms and a reminder of the flat EM-DAT classification categories was proposed to improve the user experience and facilitate linkage with existing frameworks. • The concept of unique event identifiers associated with each hazard and disaster recorded in EM-DAT, was suggested to ideally align with a unique reference across all mentions of the same event in different disaster databases.
Re-Classification and Terminology Refinements
• Several specific re-classifications and terminology refinements were suggested, such as flood subtypes, tsunamis, removal of Lahars from landslide definitions, and redefining avalanches. • A critical discussion revolved around the epidemiological data collected by EM-DAT, questioning the necessity of including 'epidemics' in the database considering the difficulties of integrating this with existing loss and impact data criteria.
Challenges and Recommendations
• The difficulty of working with variables such as 'affected' was raised, particularly in relation to reporting losses associated with disasters like drought or locust infestation. Suggestions were made to consider variables that reflect hazard aspects other than impact and intensity. • The variability of national post-disaster assessments and the lack of standardization poses a challenge for users analyzing global data. • Clarification of aggregators like "economic loss" was suggested, along with the possible need for greater granularity in the aggregated values. • The ‘impact’ information in the database could emphasize better on displaced persons, critical infrastructural damages, agricultural losses, residential versus school and hospital losses. ▪ The issue of timescale in reporting was brought up, especially in mortality counting. This led to questions about the consistency of the current tracking and recording methodology, as well as its applicability across different types of hazards. The differentiation between direct and indirect mortality may necessitate further exploration and discussion. • Participants found it challenging to reconcile EM-DAT’s current methodology and data set with future needs due to the risk of reproducing bias or incomplete data. The discrepancies between EM-DAT data and other databases for the same event were noted, and the need for clarification to users was emphasized.

Expert Presentations

The first day of the STAG meeting drew to a close with a number of brief presentations by the participating members. These presentations offered interesting insights that nicely augmented the core topics of the STAG meeting. More details or copies of the presentations can be accessed at www.emdat.be/stag.

- "The Global Crisis Data Bank, a Multi-stakeholder Initiative."
  Justin Ginnetti (IFRC).

- "Alerting the General Public to Hazards: Overview of the IFRC CAP Editor Freeware."
  Paola Yela Bello (IFRC)

- "U.S. Billion-dollar Weather and Climate Disasters: Data Tools, Multi-Hazard and Socioeconomic Risk Mapping."
  Adam Smith (NOAA)

- "Risk Information Exchange and Classifying Hazards. (RiX)"
  Adam Fysh (UNDRR)

- "New disaster damages and losses tracking system for countries."
  Iria Touzon Calle (UNDRR)

- “Extreme heatwaves in Europe 1950-2020.”
  Aglaé Jézéquel (CIRED/ LMD)

- "Disaster Displacement, Informing Action on Internal Displacement with Data, Research and Evidence."
  Sylvain Ponserre (IDMC)

- "Use of EM-DAT to Evaluate the Sendai Framework for Disaster Risk Reduction in Africa."
  Dewald Van Niekerk (African Centre for Disaster Studies) & NWU

- "UNDRR-ISC Hazard Information Profiles and their use in documenting hazards."
  Virginia Murray (UK Health Security Agency)

- "Data Requirements for Assessing Damage and Loss in Agriculture and Its Subsectors."
  Zehra Zaidi (FAO)

Wrap-up of Day 1

The first day ended with a short summary of the day's most noteworthy points. These key insights from Day 1 are summarized in Box 1.

Box 1. DAY 1 Wrap-Up and Key Messages

EM-DAT continues to serve as a crucial open-source dataset related to natural hazards and associated disaster data. The key strengths of the database, as highlighted by the STAG, include its consistency, its commitment to open access and its often overlooked yet equally significant aspects of independence and neutrality. However, the criteria for impact and loss sometimes appear less clear to users, necessitating a clear communication on aggregated values, a defined timescale, and a cutoff point for event recording and documenting.

The newly introduced disaster classification system by EM-DAT was seen as a move in the right direction. Still, the coming year can be used to further progress and finetune the system in collaboration with interested STAG members. Requests for clarification and suggestions regarding terminology were noted. These inputs might lead to further adaptations to the classification system that may be presented at the next STAG meeting.

Day 2 – 21st March 2023

Opening

The second day of the STAG meeting started with a summary of a few key highlights from Day 1 (Rhonda Stewart and Niko Speybroeck). Moving forward, it was explained that the primary objective for Day 2 would be to engage in discussions around two significant issues: addressing the problem of missing data and exploring ways to automate disaster data collection.

Plenary Session: Tackling Data Quality Issues

Presentations

Presenter: Rebecca Jones (CRED, UCLouvain)

This session was introduced by a presentation of the article "Human and economic impacts of natural disasters: can we trust the global data?" by R. Jones, D. Guha-Sapir, and Sandy Tubeuf[2]. It was explained that reliable disaster databases are essential for effective policies, but they often have missing data, resulting in an important bias. The presented study examined the global disaster database EM-DAT, identifying significant missing data for natural disasters between 1990 and 2020, particularly related to economic losses. Factors like the disaster year, income classification, and disaster type contribute to missingness. Advanced statistical methods are needed to minimize bias and ensure trustworthy global disaster data. An overview of these methods was presented.

Presenter: Damien Delforge (CRED, UCLouvain)

Damien Delforge provided insights into EM-DAT's limitations and biases, which are particularly noticeable for specific time periods, regions, or types of hazards or impacts. He subsequently highlighted potential avenues to address these challenges, structured under two major themes: i) technological opportunities and ii) data augmentation strategies.

With regard to technological opportunities, Delforge pointed to numerous technologies that could offer solutions. Among these, he highlighted satellite imagery, meteorological re-analyses, advances in natural language processing and geo-processing, and the development of application programming interfaces (APIs) for automated data flow between EM-DAT and other data collection initiatives.

Regarding data augmentation strategies, it was furthermore explained that expanding the data sources could significantly improve the robustness of EM-DAT. Delforge suggested the inclusion of demographic datasets, social media posts, online news, and data from early warning systems. As the role of artificial intelligence, specifically machine learning models, promises to play an increasing role in disaster risk reduction, ensuring high data quality is paramount. New developments offer the potential to bolster existing methodologies and retroactively strengthen datasets. This goal can be achieved through strategic partnerships, establishing communities of practice, leveraging new technologies, and maintaining active communication about progress and residual uncertainties.

Breakout Session 2

The breakout sessions that followed the presentations focused on the themes of i) handling missing data, ii) automating data collection, and iii) improving data quality, including epidemics and heatwaves. The conclusions from these breakout sessions are summed up in Table 2.

[2] https://doi.org/10.1038/s41597-022-01667-x

Table 2: Breakout Session 2 'Tackling data quality issues', Main Insights

Handling Missing Data
• A need for defining and communicating uncertainties when quantifying 'losses' was highlighted, e.g., economic loss can be quantified by infrastructure loss; but also by predicted costs of reconstruction. • There is a need to understand the discrepancies in the reported averages of economic losses across different sources. This issue was highlighted by inconsistent data for Hurricane Harvey's losses, with a difference of 45 million US dollars between various reports. While hurricanes often result in insured losses that are tracked and reported by reinsurance companies, the economic impact from uninsured floods often goes undocumented. • There are challenges in conveying missing data to relevant government agencies, particularly in data-scarce countries. Relying on local partners to encode loss and impact data when during a crisis should be done with care. • The importance of improved geo-referencing and consistency in EM-DAT for sub-national level disasters or associated events (e.g., coordinates of recorded events are often missing; some events have administrative level data, others not) was underlined. • Clear communication about changes to the system, particularly concerning geospatial data, is needed. • Encouragement for EM-DAT to mobilize universities or data producer networks for data collection and validation. • EM-DAT demonstrates the value of the human validation component. However, the human resources of the research team cannot cover the need to tap into more data sources while ensuring field validation at the sub-national level. New tools may increase data processing speed but still call for on-site observation networks. • Concerns were expressed about "data-poor locations" where information is sparse, in contrast to other parts of the world where data is more abundantly available. • The possibility of spatial modeling to help fill data gaps was mentioned. However, it was also agreed upon that modeling cannot solely compensate for historical database gaps; human verification remains essential. Likewise, relying on satellite imagery only will not suffice to validate data. • Presenting a glossary of terms detailing the taxonomy and categories used, as well as the time span issues, on the EM-DAT platform could be useful. • Handling missing data during an analysis can be done in different ways. Users can decide how to do this, being encouraged to explain the analytical strategy in a transparent manner.
Data Automation
• Automation requires a comprehensive understanding of specific parameters, including metrics to detect hazards, connect detected information to impacts, and verify sources. This understanding is vital for EM-DAT researchers in generating relevant metrics from the sources. • Despite the global proliferation of information, data gaps in disaster monitoring remain a persistent challenge (See Handling Missing Data). • The topic of application programming interfaces (APIs) was frequently a focal point during the discussions. It was suggested that EM-DAT could benefit from crafting its own API, thus promoting more effective utilization of its data. Additionally, EM-DAT might find value in leveraging APIs from external platforms, a strategy that would likely necessitate forging formal, solid partnerships to secure a sustainable data exchange process. There's a clear priority to identify parameters that would open the EM-DAT database for additional data entry while ensuring EM-DAT researchers can assess the retrieved information. APIs should be structured around the desired level of data capture and could

also be leveraged to access vast data hubs of large institutions, broadening the scope and depth of accessible information.

- EM-DAT should carefully consider API specifics, such as its type and the queries it can assist users and the EM-DAT team with. Connecting to existing APIs could result in an influx of data to EM-DAT, necessitating increased analysis capacity. It will also generate more data, which will require storage. AWS cloud-based solutions might be a viable option, but this requires careful negotiation with private sector partners like Amazon. Clarity in shared values and intent is crucial.
- APIs are built on specific metrics, facilitating communication and interaction between various stakeholders without personal contact. The goal of disseminating disaster data through APIs includes pulling data from resource-poor environments, ensuring robust existing information, and employing AI and natural language processing to expedite data processing and analysis.
- APIs are potent tools for fostering partnerships between agencies, as demonstrated by UNDRR's inter-agency experience using the World Food Program's API. The utilization of mapping polygons on GIS software, as exemplified by the work done by the Pacific Data Center, can help alleviate some data missingness.
- Scientific oversight remains critical in these processes and despite automation, maintaining these systems still necessitates further scaling up.

Improving Data Quality, Including Epidemics and Heat Waves

- Associating events with their main triggers could be beneficial in crisis-prone countries. This approach allows the highlighting of critical data for local authorities and potentially requests practical feedback through data visualization tools, despite an internal decision that is not fully in favor of producing such visualizations.
- It was deemed essential to identify phenomena with known climate drivers and differentiate them from unrelated events.
- For regions worldwide with limited data availability, the importance of local data verification networks was emphasized, especially in the context of leveraging technology for faster data processing.
- Addressing Data Quality in Epidemics
    - Regarding EM-DAT epidemic data, some STAG members underscored issues related to language, indicators, and definitions.
    - Discussions during the STAG meeting indicated the potential to develop research axes to connect natural hazards with epidemics. The need for models explaining the absence of such a link was brought up as well. A further concern that emerged was the role of 'man-made' initial disasters or disruptions that trigger epidemics. Practitioners in the field of Disaster Risk Reduction may appreciate tools that can translate the database into an analysis providing actionable insights, including man-made aspects of hazard occurrence.
- Addressing Data Quality in Heat Waves
    - Participants engaged in a dialogue about how mortality information should be linked to heatwave data. The consideration of excess mortality count data was a topic of debate.
    - Participants concurred that it could be beneficial to explore potential links between natural hazards and heat waves. This exploration requires careful modeling, especially in cases where there appears to be no obvious correlation.

Plenary Session: Priorities and Recommendations for EM-DAT

The STAG meeting concluded with a discussion on future priorities and recommendations for EM-DAT, with a continued more in-depth discussion during breakout sessions.

Breakout Session 3

The key takeaways on the 'Priorities and recommendations for EM-DAT' that resulted from the discussions during the breakout sessions are presented in Table 3.

Table 3: Breakout Session 3 'Future Priorities and Recommendations for EM-DAT', Main Insights

Data Availability and Quality
• Enhancing data availability and quality may be achieved through resources like the IFRC Global Crisis Data Bank, which leverages machine learning and may send alerts for necessary data entry. This may improve the ability to report on previously unrecorded events. • Standardizing reporting using the UNDRR classification may provide a common reference point. Translation initiatives may further increase accessibility. • The issue of missing data needs to be further discussed. E.g., even though the UNDRR's RiX will host a vast amount of data beneficial for predictive modeling, it may not be able to address the issue of missing data. • A balance between data quality and a consistent protocol should be achieved to prevent alterations with each user request. • Clear communication about data reproducibility, archived sources, and potential biases is vital. If, for example, a pre-2000 bias exists, it should be clearly mentioned to users when they access the database. The traceability of data sources should be enhanced. • The necessity of recording new types of hazards (e.g., cyber hazards) may need to be further investigated.
Reporting and Data Use Enhancements
• The implementation of reporting initiatives, such as the IFRC CAP tool, was encouraged. • The potential for supportive tools to enhance EM-DAT's usability for field actors could be explored further. • The use of APIs to facilitate partnerships is encouraged. • User interaction could be enhanced, possibly through more interactive data options, such as reports with maps and graphs, data visualization, and analysis support. • It may be worth considering the needs of secondary users, particularly those who are less visible. The underrepresentation of local authorities or NGOs in the user base highlights the importance of providing analysis-ready tools. To facilitate this, it might be essential to expand EM-DAT's ability to evaluate these requirements effectively.
Partnerships and Resource Management
• Attention may be given to helping field partners analyze EM-DAT data. • Automation will need an increase in human resources and contributions for maintenance. • Partnerships should be leveraged thoughtfully, considering both upstream and downstream of the data value chain. • Work at the sub-national level with different partners should remain in line with UNDRR codification and EM-DAT's unique identifier system. Moreover, discussions on the interoperability of current databases are needed, particularly as more data hubs and predictive modeling projects emerge.
Transparency and Communication
• Transparency in source selection and cutoff point application should be ensured. • The traceability of data sources should be improved. • Communication with priority users to determine the appropriate data set format may need to be prioritized.

Wrap-up of Day 2

The conclusion of the first day was marked by a summary of the day's most noteworthy points, followed by a courteous farewell to all participants. These key insights from Day 2 are summarized in Box 2.

Box 2: DAY 2 Wrap-up and Key Messages

Enhancing EM-DAT's Foundation: EM-DAT needs attribution of data sources, a consistent classification system, and clarity in terminology and methodology to strengthen its foundation.

Improving transparency and addressing bias: It is crucial for EM-DAT to communicate changes, methodologies, and steps taken to enhance transparency about data missingness and bias. Ongoing efforts to address bias within the EM-DAT database are essential, including expanding partnerships with institutions that collect and encode data and developing new and existing methods, such as natural language processing tools.

Promoting Collaboration Among Data Tools: With the introduction of many new data tools, it is essential to collaborate closely during the rollout phase. Learning from each other's current objectives can foster mutual growth and understanding.

Developing Meaningful Partnerships: While partnerships among data users are evolving with new technology, there's still a need for partnerships focused on systematic and reliable data collection. It's vital to ensure that people meet and communicate in all initiatives engaging with data producers.

Clarifying EM-DAT's Priorities: It is important to understand 'who we are doing this for' when reviewing EM-DAT's priorities. Identifying the needs and running the value chain, considering both upstream elements, like priority partnerships for data collection, and downstream factors, such as the capacity to provide tools and training, are critical for generating actionable feedback.

Handling Missing Data: There are multiple ways to address missing data during an analysis. It is important for users to choose a strategy that aligns with their objectives and clearly communicate this strategy to maintain transparency and comprehensibility throughout the analytic process. The matter of handling missing data requires continuous discussion and careful consideration.

Funding Diversification: During the STAG meeting, participants notably underscored the growing importance of EM-DAT, particularly in light of the increasing global initiatives and priorities. Despite the STAG report's focus on scientific advice and its omission of specific details on funding diversification or business strategies, participants concurrently expressed a clear need for increased and sustainable funding. This additional support may be vital to enhance EM-DAT's resilience and capacity in order to meet the rising demands.