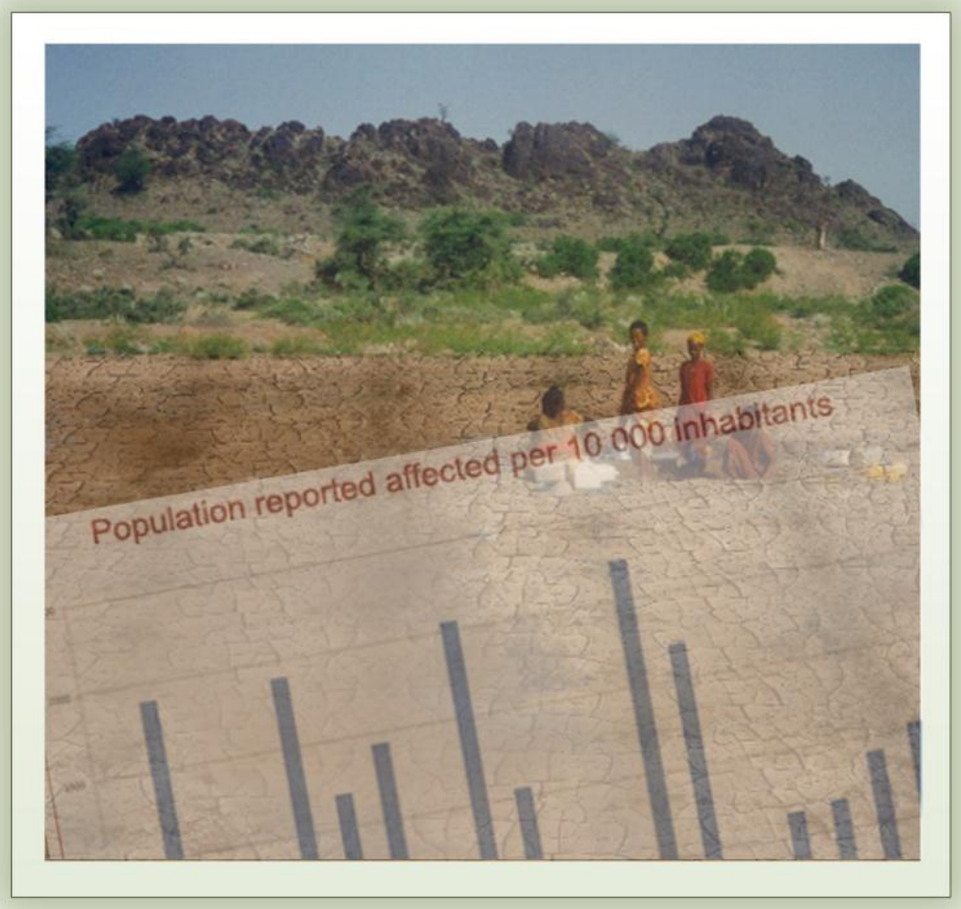


Centre for Research on the Epidemiology of Disasters
CRED

USAID
FROM THE AMERICAN PEOPLE

Moving towards Harmonization of Disaster Data: A Study of Six Asian Databases

Regina Below, Femke Vos, Debarati Guha-Sapir

CRED Working Paper No. 272

Moving towards Harmonization of Disaster Data: A Study of Six Asian Databases

Brussels, September 2010

CRED Working Paper No. 272

Regina Below
Femke Vos
Debarati Guha-Sapir

ACKOWLEDGEMENTS

This report is part of the EM-DAT Project which is funded by the US Agency for International Development's Office of Foreign Disaster Assistance (USAID/OFDA).

We would like to thank C. Villacis (UNDP/GRIP) and S. Jegillos, R.Sharma, and P. Homtong (UNDP/RCB) for their support in getting the field work accomplished in time.

We are also grateful to all the national institutions that have welcomed us in the context of this project and to the database staff in each country for their collaboration and their time.

We would like also to thank A. Wirtz (MunichRe) for her detailed comments on the report.

Finally, we would like to thank A. F. Díaz for design and layout.

Contact:

Centre for Research on the Epidemiology of Disasters (CRED)
Université catholique de Louvain
30.94 Clos Chapelle-aux-Champs
1200 Brussels, Belgium
Telephone : +32 27643327
E-mail : contact@cred.be
www.cred.be

This document is available on http://www.cred.be/sites/default/files/WP272.pdf

Printed by :
Ciaco Imprimerie, Louvain-la-Neuve (Belgium)
This publication is printed in an environmentally – friendly manner.
September, 2010

TABLE OF CONTENTS

List of tables, figures and boxes v

Executive Summary vi

List of Acronyms ix

1. Introduction and Context 1

2. Study objectives and methods 4

2.1. Goals and objectives 4
2.2. Database selection 5
2.3. Quality assessment framework and disaster database audit methodology 6

2.3.1. Quality assessment framework 7
2.3.2. Disaster database audit methodology 9

3. Case studies 11

3.1. Brief description of databases 11
3.2. Evaluation summary 16
3.3. Main finding and conclusions 24

4. Database comparison 27

4.1. Objectives 27
4.2. Comparison of presence and complementarity of data elements 27
4.3. Analysis of database completeness 35

5. Recommendations 38

6. Guidelines 42

6.1. Disaster data collection and data sources 43
6.2. Database structure 45
6.3. Data entry 46

6.3.1. Identification of the disaster events 46
6.3.2. Identification of geographical and temporal Information 49
6.3.3. Identification of the human and the economic Impact 49
6.3.4. Data Validation 50
6.3.5. Data Analysis 50
6.3.6. Database outputs and dissemination 52

7. Conclusions 54

8. Bibliography 55

ANNEX 1 : Overview of data elements in databases (comparing to EM-DAT) – stand alone databases 57
ANNEX 2 : Overview of data elements in databases (comparing to EM-DAT) – DesInventar databases 61
ANNEX 3 : Preparatory questionnaire and interview 65

Part one 65
Part two 69
Part three 81

LIST OF TABLES, FIGURES AND BOXES

Figure1: Elements of the quality framework for disaster loss database 7
Table 1: Descriptive summary of stand alone databases: Philippines, Bangladesh and Vietnam 11
Table 2: Descriptive summary of DesInventar databases: Nepal, Indonesia and Sri Lanka 13
Table 3: Evaluation of stand alone databases: Philippines, Bangladesh and Vietnam 16
Table 4: Evaluation of DesInventar databases: Nepal, Indonesia and Sri Lanka 20
Box 3.1 Conclusion from evaluation: Methodology 24
Box 3.2 Conclusion from evaluation: Accuracy and reliability 25
Box 3.3 Conclusion from evaluation: Serviceability 25
Box 3.4 Conclusion from evaluation: Accessibility 25
Box 3.5 Conclusion from evaluation: Credibility 26
Box 3.6 Conclusion from evaluation: Prerequisites and sustainability 26
Table 5: Comparison of standalone databases: Philippines, Bangladesh and Vietnam 29
Table 6: Comparison of DesInventar databases: Nepal, Indonesia and Sri Lanka 33
Table 7: Philippines - Calamidat database 35
Table 8: Bangladesh - DIDB database 35
Table 9: Vietnam - DMC database 36
Table 10: Nepal - DIMS database 36
Table 11: Indonesia - DIBI database 36
Table 12: Sri Lanka - SDIS database 37
Box 5.1 Methodology 38
Box 5.2 Accuracy and reliability 39
Box 5.3 Serviceability 39
Box 5.4 Accessibility 40
Box 5.5 Credibility 40
Box 5.6 Prerequisites and sustainability 41

EXECUTIVE SUMMARY

With disaster events increasing in magnitude and frequency, the need for disaster impact data collection and sharing is both urgent and continuous in the effort to save lives, alleviate suffering, and reduce economic losses. The systematic collection of information related to the frequency and impact of disasters is an important tool for governments, international policy setting organizations and institutions in charge of relief and recovery activities.

Sponsored by USAID/OFDA, and in the context of the Global Risk Identification Program (GRIP) of the United Nations Development Program (UNDP) framework, CRED coordinates a collaboration activity focused on national disaster data compilation initiatives. The collaboration aims at the sharing of knowledge to enhance the visibility, accessibility, and applicability of disaster databases. By providing technical support, the project contributes to strengthening the standardization, reliability, and use of existing methodological and operational approaches. A more complete and accurate collection of data on disaster occurrence and impact will ensure better risk estimations and improve the availability of information and analysis on disaster risks and risk factors. The strengthening of disaster databases will eventually serve the global, international, and national humanitarian communities involved in disaster response planning and risk reduction.

The Asian region has historically faced many challenges from disasters and remains nowadays the area most prone to natural disasters. Many initiatives aim at disaster preparedness and reduction in Asia, and among these are invaluable local disaster data compilation initiatives that support the information management of disasters at the local and international levels. In the current study, six Asian national disaster databases were selected, located in Bangladesh, Indonesia, Nepal, Philippines, Sri Lanka and Vietnam. Of these six databases, three have been developed using a standalone methodology[1], and the remaining three are based on the DesInventar methodology[2].

Based on the literature, CRED has developed a framework to capture the quality of disaster loss databases. In this framework, database quality not only is reflected in the correctness of the data but also encompasses other aspects such as database accessibility, serviceability, credibility of the database hosting institute, database methodology, and accuracy and reliability of the data. Prerequisites for maintaining a disaster database, such as the institutional environment and sustainable

[1] Standalone databases are national and sub-national databases based on a model developed by the hosting institutions: Philippines, Bangladesh, and Vietnam.
[2] Nepal, Indonesia, and Sri Lanka. See: www.desinventar.org

resources, form the basis of developing and maintaining a disaster database and are included in the framework.

A preliminary report was developed in collaboration with the 6 database hosting institutes, describing the general structure and present functioning of each disaster database. Next, case studies were performed, including on-site interviews and discussions focused on the methodological and operational procedures of the database and on identification of database strengths and weaknesses. This activity led to tailored recommendations for strengthening the databases. Next, a comparison of the global EM-DAT database and the national disaster databases was performed to study their similarities, differences, and completeness. Based on these activities, guidelines were established for the development of disaster databases and the compilation of reliable data to serve disaster data initiatives and the humanitarian community worldwide.

The effectiveness of disaster preparedness and prevention depends on the evidence base on which the programme is anchored. Equally important, disaster data is central to studies that link disasters to health, social systems, poverty or even climate change. While there is a growing recognition of the need for accurate and comparable data on the impacts of disasters, there is still much room for improvement.

Since many years, CRED has persevered in its efforts to improve the quality of data on human impacts of disasters, engaging in many methodological initiatives with collaborators. Today, with new and cheaper technologies, information on the human impact of disasters should be systematically recorded and harmonized for comparisons across regions and also against time.

Most importantly, we have realized that to design a framework to improve data quality, we needed to get down in the weeds and examine the way in which countries actually operate. We did not want to work towards a solution without a sound grasp of the issues faced by our national colleagues.

We conclude from our study that there is an urgent need for robust field methodologies to estimate the number of dead, injured, and affected and guidelines for their use by national governments, international policy setting organizations and relief agencies. Ambiguity in figures that encompass an unspecified variety of groups or conditions, a common problem in disaster impact reporting, significantly reduces the usability of the data. We strongly advocate the development of standard methodologies that every agency can use to prevent ambiguous data from becoming a source of misguided policy and erroneous decision-making.

Another key area for improvement is greater standardization of data compilation methods and definitions. This goal can be achieved only by joint international efforts to develop these tools and make them available for national-level use.

Finally, one of the limitations of our study is that we have explored only six databases with a specific context and therefore our conclusions may not necessarily apply to other scenarios. However, we feel that the lessons learnt from this exercise and from the experience of EMDAT significantly bring forward the discussion on global data harmonization and inter-operability.

LIST OF ACRONYMS

ADRC	:	Asian Disaster Reduction Center
BNPB	:	National Agency for Disaster Management
CDMP	:	Comprehensive Disaster Management Programme
CRED	:	Centre for Research on the Epidemiology of Disasters
DANA	:	Damage and Needs Assessment system
DDMFSC	:	Department of Dyke Management, Flood and Storm Control
DFID	:	Department for International Development
DIBI	:	Indonesian Disaster Information and Data
DIDB	:	Disaster Incidence Database
DIMS	:	Disaster Information/Inventory Management System
DMC	:	Disaster Management Centre
DMIC	:	Disaster Management Information Centre
DMIN	:	Disaster Management Information Network
EM-DAT	:	Emergency Events Database
EU	:	European Union
GIS	:	Geographic Information System
GRIP	:	Global Risk Identification Program
MoFDM	:	Ministry of Food and Disaster Management
NDCC	:	National Disaster Coordinating Council
NGO	:	Non Gouvernmental Organization
NSET	:	National Society for Earthquake Technology
OCD	:	Office of Civil Defense
OFDA	:	Office of Foreign Disaster Assistance
OSADI	:	Online Southeast Asia Disaster Inventory
RCB	:	Regional Center Bangkok
SDIS	:	Sri Lanka Disaster Information System
UNDP	:	United Nations Development Programme
USAID	:	United States Agency for International Development

1. INTRODUCTION AND CONTEXT

With global disaster events increasing in magnitude and frequency, the need for disaster impact data collection and sharing is both urgent and continuous in the effort to save lives, alleviate suffering, and reduce economic damage. The systematic collection of information related to the frequency and impact of disasters is an important tool for governments and institutions in charge of relief and recovery activities, as well as for the integration of disaster risk analysis and reduction.

The Asian region has historically faced many challenges from disasters and remains the area most prone to natural disasters. Several initiatives aim at disaster preparedness and reduction in Asia, and among these are invaluable local disaster data compilation initiatives that support the information management of disasters at the local and international levels.

To provide reliable disaster data, there is a need for adequate database structures, standardized methodology and operational approaches, and interoperable data formats. Improvement of disaster data analysis, as well as increased visibility of and access to disaster data, require specific focus at the smaller, intra-country special scales and on an expanded scope, by inclusion of human and economic impact indicators.

The Centre for Research on the Epidemiology of Disasters (CRED) hosts the EM-DAT database on natural and technological disasters. In EM-DAT, disaster events and their human and economic impacts are analysed at a global level. National and sub-national databases provide disaster information at smaller, intra-country scales and are complementary to the EM-DAT database.

Sponsored by USAID/OFDA, and in the context of the Global Risk Identification Program (GRIP) of the United Nations Development Program (UNDP) framework, CRED coordinates a collaboration activity focused on disaster data compilation initiatives in the Asian region. The collaboration aim is to share knowledge to improve the visibility, accessibility, and applicability of disaster databases at the national level. This goal will be achieved by helping to reinforce disaster database structures and methodological and operational approaches. Capacity building is an essential component of this activity. The strengthening of disaster databases will eventually serve the global,

international, and national humanitarian communities involved in disaster response planning and risk reduction.

This project focuses on improving capacity building at the national and regional levels, beyond that of the international EM-DAT database. This study draws from the experience of EM-DAT, a global database with commonly accepted definitions, criteria, and methodologies, but applies this experience to reinforce a more detailed and specialized effort at micro-levels. It works towards a concrete support for similar efforts, adapted for national use. This activity is therefore an audit of emerging efforts in countries, accompanied by constructive and practical help in improving and building on what already exists.

The project addresses issues of methodological and operational approach limitations that arise because of inconsistent reliability and interoperability of the data from current disaster data compilation initiatives. By providing technical support for these initiatives, the project helps strengthen the standardization, reliability, and use of existing methodological and operational approaches. Furthermore, the project will contribute to providing a more comprehensive and accurate accounting of disaster-related losses and costs to the international community. A more complete and accurate collection of data on disaster occurrence and impact will ensure better risk estimations and improve the availability of information and analysis on disaster risks and risk factors.

Because access to the actual disaster databases was beyond the control of the partners in this project and dependent on the institution compiling and housing the data, we could include six national disaster databases in our study:

- Calamidat Disaster Event Database, Philippines
- Disaster Incidence Database (DIDB), Bangladesh
- Damage and Needs Assessment System (DANA), Vietnam
- Disaster Information/Inventory Management System (DIMS), Nepal
- Indonesian Disaster Information and Data (DIBI), Indonesia
- Sri Lanka Disaster information System (SDIS), Sri Lanka

The overall aim of the present project is to provide assistance to institutions and individuals in establishing disaster data collection initiatives, allowing an expansion of the sharing of disaster data among the international community to strengthen disaster management and relief.

After an introduction describing the context of the study in Section 1, Section 2 provides an explanation of the objectives and methods used. Section 3 summarizes the findings from case studies performed in the six countries, while Section 4 presents the results from a database comparison exercise. Section 5 provides guidelines on good practices for disaster data collection initiatives for use with existing and future disaster databases worldwide.

2. STUDY, OBJECTIVES AND METHODS

Subsection 2.1 describes the goals and objectives of this study in detail. Subsection 2.2 provides an explanation of the selection of disaster databases included in the study. Finally, subsection 2.3 handles the quality assessment framework and disaster database audit methodology used in the study.

2.1. GOALS AND OBJECTIVES

The overall goal of the present project is to strengthen the quality, reliability, and sustainability of disaster databases at the (sub-)national level. This goal will be achieved by reinforcement of disaster database structures and methodological and operational approaches. The strengthening of disaster databases will eventually serve the global, international, and national humanitarian communities involved in disaster response planning and risk reduction. The objectives of the project are as follows:

- To identify disaster databases in South and Southeast Asian countries and their characteristics
- To develop a disaster database quality-assessment and audit methodology
- To identify operational and methodological strengths and weaknesses of selected databases
- To provide recommendations to reinforce disaster databases and provide remote and on-site technical assistance
- To produce general guidelines for the development of disaster databases and the compilation of reliable data

2.2. DATABASE SELECTION

Disaster data compilation initiatives in the Asian region have been identified in a previous study[1], showing one regional disaster impact database (Asian Disaster Reduction Center (ADRC)) and seventeen national disaster impact databases in Bangladesh, China, India (consisting of six sub-national databases), Indonesia, Iran, Maldives, Nepal, Philippines (two databases), Sri Lanka, Thailand, and Vietnam.

An online preliminary search of the corresponding database websites was performed to obtain an overview of the identified national disaster impact databases in the Asian region. An initial database selection for collaboration was based on the existence and accessibility of the database website, continuity of the database during the project, and relevance and number of available records in the database. National disaster databases in the following six countries were eventually selected for the study: Philippines, Bangladesh, Vietnam, Nepal, Indonesia, and Sri Lanka. Of these six national disaster databases, three have been developed based on a standalone methodology[2], and the remaining three[3] are based on the DesInventar methodology[4,5]

A preliminary report was developed in collaboration with the institutes, describing the general structure and present functioning of each disaster database. Next were on-site interviews and discussions focused on the methodological and operational procedures of the database and on identification of database strengths and weaknesses. This activity led to tailored recommendations for strengthening the databases. Next, a comparison of the global EM-DAT database and the national disaster databases was performed to study their similarities, differences, and completeness and to identify possibilities for interoperability among them. Findings of the on-site interviews and database comparisons were reported. Based on these activities, guidelines were established for the development of disaster databases and the compilation of reliable data to serve disaster data initiatives and the humanitarian community worldwide.

[1] Tschoegl L., Guha-Sapir D., Below R. An analytical review of selected data sets on natural disasters and impacts. Paper prepared for the UNDP/CRED Workshop on Improving Compilation of Reliable Data on Disaster Occurrence and Impact, Bangkok, 2–4 April 2006

[2] Standalone databases are national and sub-national databases based on a model developed by the hosting institutions: Philippines, Bangladesh, and Vietnam.

[3] Nepal, Indonesia, and Sri Lanka

[4] See: www.desinventar.org

[5] United Nations Development Programme. Guidelines and lessons for establishing and institutionalizing disaster loss databases, UNDP, 2009

2.3. QUALITY ASSESSMENT FRAMEWORK AND DISASTER DATABASE AUDIT METHODOLOGY

Worldwide, many initiatives have been started for collecting and analysing information on disasters and their associated losses. The common aims of these initiatives are to monitor disaster losses and vulnerability and to create an evidence base for allocating resources and supporting the formulation or revision of national disaster risk reduction strategies, action plans, or programmes.

Approaches to developing, structuring, and implementing a disaster database vary significantly, depending on the objectives and needs of the country, the institutional framework, and the resources available. Collecting disaster data in a database without consideration of the end goal does not carry much meaning. The importance of the data lies in its serviceability to the end user and in the outputs that arise from the data to serve user objectives. Information quality, or data quality, is thus one of the most important characteristics of a database.

According to the International Organization for Standardization (ISO), quality is defined as '*the degree to which a set of inherent characteristics fulfils the requirements*'[6]. The American National Standards Institute/American Society of Quality Control (ANSI/ASQ) defines quality as '*the totality of features and characteristics of a product or service that bears on its ability to satisfy given needs*'[7]. More specifically, data quality can be described as doing the following:
(1) consistently meeting all knowledge worker and end-customer expectations in all quality characteristics of the information products and services required to accomplish the enterprise mission (internal knowledge worker) or personal objectives (end customer); and
(2) meeting a specific degree at which information consistently meets the requirements and expectations of all knowledge workers who require it to perform their processes[8]. Many authors have published on data quality, and several frameworks for evaluating and monitoring data quality have been developed[7,8,9,10,11].

[6] David Hoyle, 2006. ISO 9000 Quality Systems Handbook. Butterworth-Heinemann.
[7] www.asq.org/glossary/q.html, last accessed on 20 September 2010.
[8] International Association for Information and Data Quality (IAIDQ) (*www.iaidq.org/main/glossary.shtml),* last accessed on 20 September 2010.
[9] FAO Statistical Data Quality Framework: A multi-layered approach to monitoring and assessment. Conference on Data Quality for International Organizations, Wiesbaden, Germany, 2004.
[10] World Bank Development Data Group and UNESCO Institute for Statistics. A framework for assessing the quality of education statistics,

In this section, a quality framework is proposed for assessing the typologies, definitions, variables, data flows, data compilation, validation, and dissemination procedures of each database. Furthermore, a disaster database audit methodology is described, and the corresponding interview questionnaire used to perform the audits is explained.

2.3.1. QUALITY ASSESSMENT FRAMEWORK

Based on the literature, CRED has developed a framework to capture the quality of disaster loss databases. In this framework, database quality not only is reflected in the correctness of the data but also encompasses other aspects such as database accessibility, serviceability, credibility of the database hosting institute, database methodology, and accuracy and reliability of the data. Prerequisites for maintaining a disaster database, such as the institutional environment and sustainable resources, form the basis of developing and maintaining a disaster database and are included in the framework. Figure 1 gives a schematic overview of these elements.

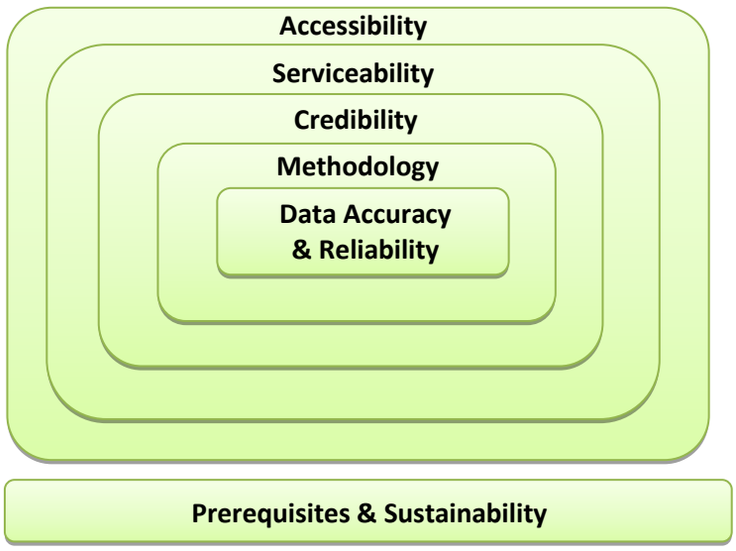


Figure 1: Elements of the quality framework for disaster loss databases

The elements make up the baseline of a qualitative assessment methodology for disaster loss databases. Each element is translated into discussion topics to build up the qualitative interview methodology. The quality framework covers all aspects

www.uis.unesco.org/TEMPLATE/pdf/SCB/DQAF%20for%20education%20statistics.pdf, last accessed on 20 September 2010

[11] International Monetary Fund (IMF) Data Quality Assessment Framework 2003, www.dsbb.imf.org/Applications/web/dqrs/dqrsdqaf/, last accessed on 20 September 2010

of the environment in which the database institute collects, inserts, analyses, and disseminates disaster loss data. Each element comprises the following main topics:

- ***Prerequisites and sustainability***

This category focuses on the institutional framework of the database institute and the level of support from the environment to the database institute for maintaining a disaster loss database. Former collaborative partnerships are described, and possibilities for future collaborations are explored. The sustainability of the database is described as the possibility of maintaining the database independently from its institutional framework, the resources and funding for the database, and the long-term objectives of the institute with regard to database continuity.

- ***Data accuracy and reliability***

This category describes the extent to which the data correctly describe the facts. First, the accuracy and reliability of the data sources and the way the institute assesses this parameter are evaluated. The focus lies on the completeness of information and geographical coverage of data sources, as well as on possible biases in using data sources. Second, the extent of the accuracy and reliability of the data present in the database is evaluated based on the level of standardization of data entry procedures, selection and number of data sources, and data validation procedures. The minimum required information for entering and analysing data is assessed. The availability of guidelines and training possibilities for database staff are also investigated.

- ***Methodology***

This category focuses on the application of clear and sound concepts and definitions. The entry criteria for the inclusion of events in the database as well as disaster definitions and classifications are described. The practice of data collection and data provision from data sources are investigated. The extent to which entered data correspond to the objectives of the database is described by the type of information entered, the frequency of data entry, and the type of database used. The procedures for analysing data extracted from the database are evaluated, and the internal and external use of the analytical products is discussed. Ways of storing and backing up the data are described.

- ***Credibility***

Credibility refers to the institute's established expertise within the domain, its impartiality, and its transparency. The institute's involvement in developing professionalism is evaluated, and the awareness and assurance of data quality by the management is discussed. The availability of information on a website or other medium concerning the institute, database goals and objectives, methodology,

data dissemination, and changes in practice and policies regarding the database are evaluated.

- ***Serviceability***

This category describes the usefulness and convenience of the presented information and the assistance provided to the users when using the database. The focus is on the timeliness, also called periodicity, of data dissemination. The availability of user documentation is investigated, as is information about the methodology that is necessary for understanding and making a correct interpretation and use of data, outputs, and analytical functions. The ease of interpretation of the information, the user profiles, and perceived relevance of the information by the users are also addressed. The level of interoperability of the database is evaluated by focusing on participation in networks sharing common objectives, data formats, and possibilities for data exchange, and the use of common identifiers, standards, and classifications.

- ***Accessibility***

This category focuses on the policy applied by the institute concerning data access and restrictions and the availability of contact details in case of further information demand.

2.3.2. DISASTER DATABASE AUDIT METHODOLOGY

The methodology aims at identifying the operational and methodological weaknesses that hinder the interoperability, reliability, and use of disaster databases. Issues are addressed such as data flows from sub-national administrative levels and their aggregation. The audits form the basis for tailored reports for each selected database. Each report includes strengths and weaknesses and recommendations to strengthen the quality and reliability of the data and the sustainability of the database. These reports help to support data analysis, collection, and processing.

- ***Gathering of disaster loss data compilation initiatives***

The first step in an assessment of disaster loss database quality and reliability was to contact the existing initiatives. Database managers and directors from different Asian institutes were contacted by email to inform them about the current study and to invite them to participate. In parallel, key respondents were contacted for more information about disaster loss data compilation initiatives in the Asian region.

- ***Preliminary (exploratory) questionnaire***

To target the survey and become familiar with the context of the selected disaster loss data compilation initiatives, general information was compiled, starting with a description of the database objectives and purposes, database denomination, information about the institute in charge of hosting the database, and contact details for the database contact person. Next, the contents of the online datasets were reviewed. The disaster types and definitions were described, as well as human and economic impact indicators. Information was compiled on the coverage of the data in space and time, the level of observation and resolution, the present state of the initiative, and the most recent updates of the data. The presence or absence of a disaster identification code and the total number of events in the database (both online and offline) were assessed. If a website was available for accessing the data, the language and content of the website were described, as well as the available outputs.

- ***Audits***

Survey research is the method of gathering data from respondents thought to be representative of a specific population, using an instrument composed of closed, structured, or open-ended questions. Whereas quantitative research responds to research questions based on statistical projections, qualitative research responds to research questions by thorough investigation and understanding of the subjects of interest. In the present study, the research tools were a qualitative questionnaire and personal interviews.

Because the project was designed as an exploratory study and meant to investigate specific contexts, the primary study tool selected was the in-depth interview. A structured interview was developed that combined open questions, filter (contingency) questions, multi-option questions, and dichotomous questions. During development of the interview questionnaire, attention was paid to the comprehensiveness of the questions, their relevance, the level of detail and specificity, the logical order and context of the questions, and the prevention of biases and influencing questions. The interviews were structured around the main elements of the quality framework. Annex 3 shows the interview categories and topics within each category and shows the detailed interview questionnaire used to study the quality and reliability of disaster loss data-compilation initiatives.

3. CASE STUDIES

The following section is divided into three subsections aimed at presenting and summarizing the results of the interviews conducted in the six Asian countries. Subsection 3.1 gives a descriptive summary of each database, and subsection 3.2 presents a summary of the evaluation of each database. Finally, subsection 3.3 displays the overall results of the interviews.

3.1. BRIEF DESCRIPTION OF DATABASES

The following tables give a short description and information about each database, based on part A of the questionnaire (see Annex 3). For increased clarity and consistency, the databases are separated into two categories: databases with a standalone methodology (Philippines, Bangladesh, and Vietnam; Table 1) and databases using the DesInventar methodology (Nepal, Indonesia, and Sri Lanka; Table 2).

Table 1: Descriptive summary of stand alone databases: Philippines, Bangladesh and Vietnam

PHILIPPINES	
Name and acronym	Calamidat Disaster Event Database (Calamidat)
Managing organization	Office of Civil Defense (OCD), National Disaster Coordinating Council (NDCC)
Database objective	Enhance capacity for disaster analysis and decision support for disaster risk management, Institutionalize GLIDE compliant disaster event database system within country
Short description	CALAMIDAT.PH is an internet-based, GLIDE associated national disaster event database system that serves as a tool to support evidence-based preparedness and mitigation initiatives for disaster risk management
Online database (URL)	http://www.calamidat.ph/dm/web/[12]
Level of observation	Sub-national
Level of resolution	National and local
Geographical coverage	National
Disasters	Natural disasters, Human-induced disasters, Complex emergencies
Time coverage	1969-2009 (42 years)

[12] At the time of reporting, the database is under development and not yet accessible.

Total entries	590
Indicators (human impact)	Deaths, missing, injured, affected, displaced, casualty, internally displaced persons, victims, survivors
Indicators (other)	Economic losses, Houses, infrastructure, agriculture, fisheries, schools, private damaged/destroyed
Language	English
BANGLADESH	
Name and acronym	Disaster Incidence Database (DIDB)
Management organization	Comprehensive Disaster Management Programme (CDMP), Ministry of Food and Disaster Management (MoFDM)
Database objective	Track disaster event and store relevant information on disasters in Bangladesh
Short description	GIS-based open source database; its content is focused on recent disaster events. Interactive web-based system consisting of a tabbed interface that includes tables, dynamic query and maps.
Online database(URL)	www.dmic.org.bd/didb[13]
Level of observation	National and Sub-national
Level of resolution	Sub-districts
Geographical coverage	National
Disasters coverage	Natural disasters, Technological Disasters, Complex Emergencies
Time coverage	1970-2009 (30 years)
Total entries	76
Indicators (human impact)	Fatalities, missing, injured, affected, evacuated
Indicators (other)	Economic damage, sector damage, infrastructure damage, aid contribution
Entry criteria	None
Language	English
VIETNAM	
Name and acronym	Damage and Needs Assessment system (DANA)
Management organization	Department of Dyke Management, Flood and Storm Control (DDMFSC); Disaster Management Centre (DMC)
Database objective	Identify severity and extent of negative impacts of disasters on human life, economy and environment in the disaster prone areas; thereby proposing options for

[13] At the time of reporting, the database is not yet accessible and is still under development

	rehabilitation and recovery.
Short description	Online database with information on natural disasters by event; disaggregated at provincial level, as well as yearly summaries per type of disaster at national level.
Online database (URL)	http://www.ccfsc.org.vn/KW6F2B34/Disaster-Database.aspx
Level of observation	Local government
Level of resolution	Provincial
Geographical coverage	National
Disasters coverage	Natural water-related disasters
Time coverage	1989-2008 (20 years)
Total entries	211
Indicators (human impact)	Deaths, missing, injured, affected, people lost all property, people needing aid
Indicators (other)	Economic losses, housing, school, hospital, agriculture, irrigation, transportation, fisheries, communication, energy
Language	English – Vietnamese

Table 2: Descriptive summary of DesInventar databases: Nepal, Indonesia and Sri Lanka

NEPAL	
Name and acronym	Disaster Information/Inventory Management System (DIMS)
Management organization	National Society for Earthquake Technology (NSET)
Database objective	Establish comprehensive and consistent data inventory system of disasters to support Government and serve different levels of Government authorities for decision making, as well as NGO's for project implementation, and for awareness raising of the general public.
Short description	Database available on-line through the DesInventar website. Interactive web-based system consisting of a tabbed interface, including query, data, map, chart, statistics and reports options.
Online database(URL)	http://www.desinventar.net/
Level of observation	District
Level of resolution	Village Development Community
Geographical coverage	National

Disasters coverage	All natural disasters, Human-induced disasters
Time coverage	1971 - 2003 (38 years)[14]
Total entries	16,879
Indicators (human impact)	Deaths, missing, injured, affected, evacuated, relocated, victims
Indicators (other)	Economic losses, houses, routes, farming/forest, livestock, education centers, medical centers, transport, agriculture, communications, power, education, relief, water supply, sewage and drainage, industry, health
Language	English
INDONESIA	
Name and acronym	Indonesian Disaster Information and Data (DIBI)
Management Organization	National Agency for Disaster Management (BNPB)
Database objective	Provide data for risk identification, policy formulation and decision making, ultimately ensuring that funds are channelled to risk reduction based on the trends and patterns identified through Dibi-based analysis
Short description	Interactive web-based system consisting of a tabbed interface, including pre-made summary tables, query, data, map, chart, statistics and reports options
Online database (URL)	http://dibi.bnpb.go.id; http://www.desinventar.net/
Level of observation	Provincial
Level of resolution	District
Geographical coverage	National
Disasters coverage	Natural disasters + non-natural disaster, social disasters
Time coverage	1997-2009 (13 years)
Total entries	6,110
Indicators (human impact)	Deaths, missing, injured, affected, evacuated
Indicators (other)	Economic losses, houses, health facilities, education facilities, rice fields, roads, worship facilities, offices, kiosks, infrastructure, plantations, ponds
Language	English, Bahasa
SRI LANKA	
Name and acronym	Sri Lanka Disaster information System (SDIS)
Management organization	Disaster Management Center (DMC)

[14] Data collection has been ongoing from 2003 to the present; however, data from 2008 onwards are not publicly available.

Database objective	Input as vulnerability layer for risk assessment models ('proxy' indicators); Support for planning (Preparedness, risk mitigation); Follow-up of efficiency of these plans, Validation of risk & hazard maps; Support for policies/regulations and investments; Damage assessment system in major disasters
Short description	Interactive, web-based system consisting of a tabbed interface that includes tables, dynamic queries, data, map, chart, statistics, exportable data option and reports options
Online database (URL)	http://www.desinventar.lk
Level of observation	Sub-national
Level of resolution	Local: (Secretariat Division)
Geographical coverage	Nationwide
Disasters coverage	Natural disaster, technological disasters
Time coverage	1974 – 2009 (36 years)
Total entries	100,846
Indicators (human impact)	Deaths, missing, injured, affected, relocated, victims, evacuated
Indicators (other)	Economic losses, housing, crops, cattle, education centers, hospitals, transportation, agriculture, power and energy, communications, water supply, industries, relief, sewage, education, health sector
Language	English

3.2. EVALUATION SUMMARY

The following section provides a summary of the evaluation that was based on part B of the questionnaire (see Annex 3). It is delineated by the following six topics: methodology, accuracy and reliability, serviceability, accessibility, credibility, and prerequisites and sustainability. Table 3 shows the databases with a standalone methodology (Philippines, Bangladesh, and Vietnam), and Table 4 shows databases using the DesInventar methodology (Nepal, Indonesia, and Sri Lanka).

Table 3: Evaluation summary of standalone databases: Philippines, Bangladesh and Vietnam

Philippines, Calamidat Disaster Event Database	
METHODOLOGY	***Concepts and definitions***: Use of national standards ***Entry criteria***: Destructive cyclones, all disasters with significant effects ***Disaster classification***: Equal level ***Data collection***: Standardized collection form ***Data entry***: Standardized entry form; data control at different levels ***Data analysis***: Basic, through the integrated system ***Disaster identification number***: Yes ***Database system***: MySQL (relational database), no geocoding component; analytical integrated system
ACCURACY AND RELIABILITY	***Data sources****:* Government + Complementary sources (e.g. Philippine Coast Guard) ***Database****:* Standard compilation procedure; Cross-checking and validation procedures ***Training****:* Staff supervision; No specific training; No written guidelines – verbal training of staff
SERVICEABILITY	***Outputs and functions***: Online querying information; Limited analytical tools and access to full event reports (PDF format); ***Interpretability***: Limited information ***Timeliness/periodicity***: Data available within 2-3 weeks ***Interoperability***: Use of GLIDE number; Part of a network sharing common objectives (ADRC and OSADI) ***User documentation***: In preparation ***User profile***: None ***Main users***: Researchers, students, local officials, international organizations, NGO's and NDCC member agencies

ACCESSIBILITY	***Access to the database***: No cost and restrictions to access to data (login needed) ***Contact details***: Yes
CREDIBILITY	***Transparency****:* Information available on the website ***Expertise****:* Report from OCD and NDCC member agencies; data used when new events occur and during press conferences *Quality management:* Management body supportive of quality improvement ***Impartiality****:* Database indirectly used for national resource allocations
PREREQUISITES AND SUSTAINABILITY	***Institutional framework****:* Government; stated in the OCD mandate ***Resources****:* ADRC ***Collaboration network****:* ADRC, OSADI/ASEAN, NDCC member agencies ***Continuity/Long-term objectives****:* Joint collaboration to produce analytical products and improve interoperability
Bangladesh, Disaster Incidence Database	
METHODOLOGY	***Concept and definition****:* Standard definition ***Entry criteria****:* None ***Disaster Classification****:* Equal level ***Data collection****:* Standardized collection form ***Data entry form****:* Standardized entry form ***Data analysis****:* Not integrated in the system, only a GIS component ***Disaster identification number****:* Yes ***Database system****:* PostgreSQL and PostGIS
ACCURACY AND RELIABILITY	***Data sources****:* Governmental (3); NGO's; Private and governmental press; priority given to Gov. sources ***Database****:* Standard compilation procedure (limited at event and national level) ***Training****:* No specific training; No written guideline

SERVICEABILITY	***Outputs and functions***: Online querying information limited to reports (PDF) + GIS function (maps); No specific outputs ***Interpretability***: Limited functions ***Timeliness/periodicity***: +/- real time ***Interoperability***: Not part of a network ; Use of GLIDE ***User documentation***: In preparation ***User profile***: None ***Main users***: Governmental agencies, researchers
ACCESSIBILITY	***Access to the database***: No cost and restrictions but access limited ***Contact details***: Yes
CREDIBILITY	***Transparency:*** No information at this stage (database not used widely) ***Expertise:*** No reports ***Quality management:*** Management body supportive of quality improvement ***Impartiality:*** Database used for national resource allocations
PREREQUISITES AND SUSTAINABILITY	***Institutional framework:*** Government ***Resources:*** UNDP, EU and DFID ***Collaboration network:*** None ***Continuity:*** Database is in its early stage of development; Additional information at sub-district levels; Keep ongoing with more comprehensive approach and collaboration with other databases network
Vietnam, Damage and Needs Assessment system	
METHODOLOGY	***Concept and definitions:*** Standard definitions ***Entry criteria:*** None ***Disaster classification:*** Equal level ***Data collection:*** Standardized collection form ***Data entry:*** Standardized entry form ***Data analysis:*** None ***Disaster identification number:*** Yes ***Database system***: Flat database linked to Excel and PDF files

ACCURACY AND RELIABILITY	***Data sources:*** Governmental (Central Committee for Flood and Strom Control (CCFSC)) ***Database:*** Standard compilation procedure, data checked with data source ***Training:*** No specific training; Guideline available
SERVICEABILITY	***Outputs and functions***: Online querying information; access to full event reports (PDF) and yearly summary (i.e. static tables); no GIS function; No specific outputs ***Interpretability***: No supporting documents ***Timeliness/periodicity***: 2 weeks ***Interoperability***: Not part of a network ; No use of GLIDE ***User documentation***: Guidelines available (only in Vietnamese language) ***User profile***: Not monitored ***Main users***: Government; All stakeholders in disaster management
ACCESSIBILITY	***Access to the database***: No cost or restrictions to access data ***Contact details***: Available
CREDIBILITY	***Transparency:*** Limited information available on website ***Expertise:*** No reports ***Quality management:*** Management body supportive of quality improvement ***Impartiality:*** Database used for national resource allocations
PREREQUISITES AND SUSTAINABILITY	***Institutional framework:*** Government ***Resources:*** Sustainable funding is a major hindering factor for the maintenance and development of the database ***Collaboration network:*** Donors ***Continuity:*** Long-term objective of implementing a relational web-based database system and monitoring at local and national level

Table 4: Evaluation summary of DesInventar databases: Nepal, Indonesia and Sri Lanka

Nepal, Disaster Information/Inventory Management System	
METHODOLOGY	***Concept and definitions**: Use of standard definitions* ***Entry criteria**: None* ***Disaster classification**: Hierarchical* ***Data collection**: Standardized collection form* ***Data entry**: Standardized entry form* ***Data analysis**: Through DesConsultar software and separate analyses in excel* ***Disaster identification number**: Yes* ***Database system**: SQL database*
ACCURACY AND RELIABILITY	***Data sources**: National newspapers, government; Priority given to the media; No source checking* ***Database**: Standard compilation procedure; No standard validation procedures but random checking* ***Training**: Initial training, lack of follow-up*
SERVICEABILITY	***Outputs and functions**: Online querying and data extraction; Analytical reports* ***Interpretability**: No supporting documents* ***Timeliness/periodicity**: Yearly basis* ***Interoperability**: No use of GLIDE; Common format of the database shared with all DesInventar databases; No data sharing* ***User documentation**: Guidelines available* ***User profile**: Not monitored* ***Main users**: Government; (I)NGO's; wider community*
ACCESSIBILITY	***Access to the database**: No cost and restrictions to access to data (accessible from 1971 to 2007)* ***Contact details**: Available*

CREDIBILITY	***Transparency**: Information available on website* ***Expertise**: database-related conferences and papers* ***Quality management**: Management supportive of quality improvement* ***Impartiality**: Database not used for resource allocations*
PREREQUISITES AND SUSTAINABILITY	***Institutional framework**: Basic functions of database maintained independently* ***Resources**: Lacking (except data entry)* ***Collaboration network**: None* ***Continuity**: Long-term objective of institutionalizing disaster inventory at local and national level*
Indonesia, Indonesian Disaster Information and Data	
METHODOLOGY	***Concept and definitions**: Use of standard definitions* ***Entry criteria**: None* ***Disaster classification**: Equal* ***Data collection**: Standardized collection form* ***Data entry**: Standardized entry form* ***Data analysis**: Through DesConsultar software and separate analysis in Excel and ArcGis* ***Disaster identification number**: Yes* ***Database system**: PostGreSQL in Linux environment*
ACCURACY AND RELIABILITY	***Data sources**: Local governments validated by national Government; University* ***Database**: Validation procedure, random checks once entered in **database*** ***Training**: Training for users and database administrators*

SERVICEABILITY	***Outputs and functions**: Online querying and data extraction; Printed analytical reviews; Maps and hazard prone indices* ***Interpretability**: No supporting documents* ***Timeliness/periodicity:** Yearly basis* ***Interoperability**: No use of GLIDE; Common format of the database shared with all DesInventar databases; Data sharing within the country (e.g.SIMPADU)* ***User documentation:** Available* ***User profile:** Monitoring of user profiles* ***Main users:** BNPB staff, ministries and sub-national administrations, all stakeholders in DRR*
ACCESSIBILITY	***Access to the database:** No cost and restrictions to access data (accessible from 1997 to 2008); non-published data available on request* ***Contact details:** Available*
CREDIBILITY	***Transparency:** Information available in different reports and on website* ***Expertise:** Conference attended* ***Quality management:** Management supportive of quality improvement* ***Impartiality:** Database used for national resource allocations; Role in respectability*
PREREQUISITES AND SUSTAINABILITY	***Institutional framework:** Government, with support from UNDPResources: SCDRR-Indonesia, UNDP, Government* ***Collaboration network:** UNDP/BCPR, UNDP-Indonesia , La Red, SIMPADU PNPM Mandiri, Data and Information Forum, National institutes* ***Continuity:** Long-term objective to create sub-national database platforms; build capacity for database management at local level; illustrate costs of disasters in terms of losses in the development progress*

Sri Lanka, Sri Lanka Disaster information System	
METHODOLOGY	***Concept and definition**: Use of standard definition* ***Entry criteria**: None* ***Disaster classification**: Equal* ***Data collection**: Standardized collection form* ***Data entry**: Standardized entry form* ***Data analysis**: Through DesConsultar software and separate analysis in Excel* ***Disaster identification number**: No* ***Database system**: SQL database*
ACCURACY AND RELIABILITY	***Data sources**: Governmental sources and press (historical data); Priority given to Gov .sources;* ***Database**: Standard compilation procedure; Cross-checking and validation procedures* ***Training**: Depends on funding, user guideline integrated into DesInventar but –additional training is required*
SERVICEABILITY	***Outputs and functions**: Online querying information and data extraction ; Preliminary report published in 2007; New report on disaster risk poverty and human development + district level based report under progress* ***Interpretability**: Through DesConsultar software* ***Timeliness/periodicity**: Daily* ***Interoperability**: No use of GLIDE; Common format of the database shared with all DesInventar databases; Project of linking with other databases within the countries* ***User documentation:** Guidelines integrated into the system, but not sufficient; a minimum level of knowledge and computer skills required* ***User profile**: No monitoring of users profile but DMC and UNDP working closely together to address user needs* ***Main users**: Governmental agencies, NGO's, researchers and students, media and technical agencies*
ACCESSIBILITY	***Access to the database**: No cost and restrictions to access to data* ***Contact details**: Minimum contact details (through DMC website)*

CREDIBILITY	***Transparency**: Information available* ***Expertise**: Information on data analysis and dissemination* ***Quality management**: Management body supportive of quality improvement* ***Impartiality**: Database used for national resource allocations*
PREREQUISITES AND SUSTAINABILITY	***Institutional framework**: Government with the support of UNDP and LaRED* ***Resources**: UNDP/Regional Centre Bangkok and the UNDP country office* ***Collaboration network**: UNDP, DesInventar/LaRED, Governmental agencies* ***Continuity**: Linking the database with other datasets from Governmental agencies in order to complete the information*

3.3. MAIN FINDINGS AND CONCLUSION

The following section presents overall findings based on the database audits. The section is structured according to six main topics: methodology, accuracy and reliability, serviceability, accessibility, credibility, and prerequisites and sustainability, emphasizing for each topic the strengths and weaknesses that were identified in the six disaster databases.

Box 3.1 Conclusion from evaluation: Methodology

Strengths

- Standard methodology applied since the beginning
- Historical data existing for at least 10 years to measure disaster trends
- Standard definitions used for disaster types
- Standard form used for collecting information
- Standard template used for data entry

Weaknesses

- Absence of entry criteria or impact threshold for entering disaster events in most of the databases
- Lack of use of internationally recognized standard definitions for disaster events and variables
- Lack of use of a disaster event classification or hierarchical classification
- Lack of commonly accepted and standardized methodology to collect economic loss data
- Lack of use of a unique ID number in most of the databases
- Definition of impact indicators not always complete

Box 3.2 Conclusion from evaluation: Accuracy and reliability

Strengths
- Country-wide data coverage
- All main human indicators (deaths, missing, injured, and affected) included in database

Weaknesses
- Priority mainly given to governmental sources (media are usually used as a secondary source), which can lead to accuracy questions
- Validation process executed but often limited
- Missing values/data

Box 3.3 Conclusion from evaluation: Serviceability

Strengths
- Database accessible in English
- Database still functional and updated at regular intervals

Weaknesses
- Limited supporting material for interpretation of outputs
- User documentation often incomplete or not clearly stated, which may lead to inappropriate use of data or misunderstanding
- Collaboration network limited, organization often serves country priorities and objectives
- Lack of knowledge about the end uses and users
- Analytical capacities usually limited, thus a need for outputs
- Development of querying functions, search tools need to be developed

Box 3.4 Conclusion from evaluation: Accessibility

Strengths
- Complete database available online
- Free of charge
- No restrictions to access data; only technical barriers

Weaknesses
- Dissemination of the information usually at a country level

Box 3.5 Conclusion from evaluation: Credibility

Strengths

- The management body usually supportive of quality improvement of the database over the longer term

Weaknesses

- Information on the database (goals, objectives, methodology, outputs, etc.) often limited or not clearly stated (transparency)

Box 3.6 Conclusion from evaluation: Prerequisites and sustainability

Strengths

- Database recognized as a reference within the country
- Sustainability ensured even if databases not developed or hosted within a governmental institute
- Awareness of need for disaster impact database

Weaknesses

- Additional staff/funding needed, mostly for development of outputs, analytical products, integrated systems

4. DATABASE COMPARISON

This section describes the summary findings from a comparative exercise between the audited national databases and the EM-DAT database. Section 4.1 provides an explanation of the objectives of the comparative exercise. Section 4.2 gives a comparison of presence and complementarity of data elements, and section 4.3 provides an analysis of database completeness.

4.1. OBJECTIVES

The comparative exercise involves comparison of database structures and methodologies from the CRED EM-DAT international disaster database and the national disaster databases. The objectives of the present analysis are as follows:

- To study the similarities and differences between database structures
- To study the completeness of reported information in both databases
- To report on the complementarity and possibilities for interoperability between databases

The aim of the exercise is to generate knowledge on database completeness, accuracy, and interoperability.

4.2. COMPARISON OF PRESENCE AND COMPLEMENTARY OF DATA ELEMENTS

Data elements are the field names of the indicators that build up the data sheet in the database. Examples of field names are serial identification number, disaster type, geographical information, date of occurrence, human impact, and economic impact.

Basic data elements are those fields that are indispensible for the logic and structure of the database and for the ability to perform useful analysis based on the recorded disaster data. Basic data elements in disaster databases are event identification code, disaster type, geographical location, start and end date of disaster occurrence, human impact, and economic or structural impact. Generally

speaking, besides the presence or absence of these basic data elements, the classification, definitions, and standards applied within the database are important features of a useful and user-friendly disaster database.

The comparison of the presence of data elements is based on the database structure and the raw data from both databases as they are being used in practice. This means that 'missing' data elements include both data elements that are not present in the database as well as data elements present in the database that do not include any value (only zeros or blanks). The Annexes 1 and 2 provides an overview of the data elements that are missing or present in both databases.

When information is compiled in one database but lacking in another, exchange of information could strengthen a database by making it more comprehensive. The complementarity of data elements is investigated between the databases, based on the outcomes in section 1.1. Tables 5 and 6 summarize the results of this exercise.

Table 5: Comparison of standalone databases: Philippines, Bangladesh and Vietnam

Philippines (Calamidat)	Bangladesh (DIDB)	Vietnam (DMC)
The present structure of the database provides a searchable list of disaster events in Philippines, from which each event can be downloaded as a PDF file. However, the access to the data should be facilitated (i.e., access to an Excel sheet) and additional functionalities must be developed. The disaster events are identified by a disaster identification number as well as with the GLIDE number. The disaster data are presented on the website at a national level; further information (i.e., breakdown at the sub-administrative level) is available in the NDCC situation report only. Disasters in both databases can be more easily linked as they are aggregated at the national level.	The present structure of the database is quite simple and must be developed to facilitate the entry of and access to the data. Data cannot be imported, and figures are presented in the database in a summary text, but detailed reports can be downloaded in a PDF format. The disaster events are identified by a disaster identification number as well as with the GLIDE number. The disaster data are presented on the website at a national level; further information (i.e., breakdown at the sub-administrative level) is available in the situation report only on the Disaster Management Information Network (DMIN) portal. Disasters in both databases can be more easily linked as they are aggregated at the national level.	The present structure of the database provides a searchable list of disaster events in Vietnam, presented as downloadable PDF files. The disasters are not identified by a disaster identification code. Data can be retrieved either per event showing information at the province level, or as a countrywide annual summary. Disasters in both databases are more easily linked if a disaster identification code is used, thereby increasing the interoperability between the databases.
Associated disasters are disasters triggered by a primary disaster, for example, a landslide that is triggered by an earthquake. A disaster database should ideally store data on the full human impact of the primary and associated disasters; however, in practice, it can be complicated to measure the impacts of a secondary event apart from its triggering event. A disaster classification that distinguishes primary and associated events could contribute to providing clear statistics on disaster impacts.		
The duration of a disaster event is an important feature that determines the full impact of a disaster. In parallel with the duration (start date – end date) of the disaster itself, the duration of the disaster impact is questioned. For example, when do the homeless people have shelter again? Measurement of the duration of the disaster and its impact should be clearly defined and consistently applied within the database. Providing information on start and end dates in the database also facilitates comparability between datasets.		

Philippines (Calamidat)	Bangladesh (DIDB)	Vietnam (DMC)
The disaster area should ideally be represented by the disaster footprint, i.e., a polygon, or the coordinates of the affected cities. Although recent efforts have addressed the issues of attributing disaster footprints to disaster location, currently there is no standardized method used by international and national databases. Neither the EM-DAT nor the Calamidat database includes coordinates on the disaster impact. The storage of coordinates, or ideally disaster footprints, could benefit the compatibility between databases and could be a basis for enhanced outputs for policy makers and the wider humanitarian community. Adding geocoding information remains a priority as OCD is working closely with the Online Southeast Asia Disaster Inventory (OSADI), which has also a geocoding component.	The DIDB was first developed as a GIS-based open-source database directly connected to a GIS interface component. Therefore, all geocoded information is included in the database. The storage of coordinates, or ideally disaster footprints, could benefit the compatibility between databases and could be a basis for enhanced outputs for policy makers and the wider humanitarian community.	The disaster area should ideally be represented by the disaster footprint, i.e., a polygon, or the coordinates of the affected cities. Although recent efforts have addressed the issues of attributing disaster footprints to disaster location, currently there is no standardized method used by international and national databases. Neither the EM-DAT nor the DMC database includes coordinates on the disaster impact. The storage of coordinates, or ideally disaster footprints, could benefit the compatibility between databases and could be a basis for enhanced outputs for policy makers and the wider humanitarian community.

Philippines (Calamidat)	Bangladesh (DIDB)	Vietnam (DMC)
Standardized human and economic impact definitions are the basis for consistent and compatible databases. In EM-DAT, the number of people killed is defined as 'Persons confirmed as dead and persons missing and presumed dead', and includes as such missing people. Calamidat defines deaths and missing people as separate indicators. The incorporation of a data element 'Number of people missing' in the EM-DAT database would facilitate interoperability between the databases. In addition, the breakdown information for the economic losses in Calamidat is quite complete, and the other databases should take the benefit of their experience and methodology in assessing the economic impact to have standard methods.	Standardized human and economic impact definitions are the basis for consistent and compatible databases. In EM-DAT, the number of people killed is defined as 'Persons confirmed as dead and persons missing and presumed dead', and includes as such missing people. Calamidat defines deaths and missing people as separate indicators. The incorporation of a data element 'Number of people missing' in the EM-DAT database would facilitate interoperability between the databases.	Standardized human and economic impact definitions are the basis for consistent and compatible databases. In EM-DAT, the number of people killed is defined as 'Persons confirmed as dead and persons missing and presumed dead', and includes as such missing people. DMC defines killed persons as 'people whose deaths are caused by disaster's direct impact with their found bodies' and missing people as 'people who are living in the reporting administrative unit but dead by disaster without finding their bodies or any information after disaster'. The incorporation of a data element 'Number of people missing' in the EM-DAT database would facilitate interoperability between the databases.
The characteristics of the disaster, such as the magnitude, event name, and the local time, can support the understanding of a disaster that has struck a population. Because Calamidat monitors major events, it could complement the information stored in the EM-DAT database.	Characteristics of the disaster, such as the magnitude, event name, and the local time, can support understanding of a disaster that has struck a population. EM-DAT provides these characteristics and could complement the information stored in both the DIDB and DMC databases.	

Philippines (Calamidat)	Bangladesh (DIDB)	Vietnam (DMC)
The Calamidat database provides detailed and comprehensive information on the physical impacts of disasters. Examples are the number of houses damaged or destroyed. This information is provided per disaster at the lowest administrative level (barnaguays). Although the current database structure does not allow for a rapid and easy exchange of data, the information stored in PDF files could complement the information on physical impacts in EM-DAT.	The DIDB database provides detailed and comprehensive information on the physical impacts of disasters. Examples are the number of houses partially or fully destroyed, area of damaged crops, number of livestock deaths, educational institute partially or fully affected, km of roads partially or fully damaged as well as bridge/culvert and km embankments damaged. This information is provided per disaster at the district level. Although the current database structure does not allow for a rapid and easy exchange of data, the information stored in PDF files could complement the information on physical impacts in EM-DAT.	The DMC database provides detailed and comprehensive information on the physical impacts of disasters. Examples are the number of houses collapsed or washed away, area of rice fields flooded, length of dyke eroded, number of bridges and sluices damaged, and amount of electric wire damaged. This information is provided per disaster at the province level. Although the current database structure does not allow for a rapid and easy exchange of data, the information stored in PDF files could complement the information on physical impacts in EM-DAT.
The names of the affected regions, provinces, and districts provided in the three databases can provide complementary information to increase the resolution of EM-DAT's national-level figures.		

Table 6: Comparision of DesInventar Databases: Nepal, Indonesia and Sri Lanka

Nepal (DIMS)	Indonesia (DIBI)	Sri Lanka (SDIS)
In none of these databases is each entry linked to a serial identification code, a region code, a district code, and a village code. A unique disaster identification number per event is not included in the database. As such, disaster events are recorded per administrative unit and not per event. In the EM-DAT database, disasters are recorded with a unique identification number per event (first level = event level). Disaster impact data are recorded at the country level (second level). Data can be retrieved either per event or per country. Disasters in both databases could be easily linked if an event-level disaster identification code is used, increasing the interoperability between the databases.		
Associated disasters are disasters that are triggered by a primary disaster; for example, a landslide that is triggered by an earthquake. The database should store data on the full human impact of a disaster; however, in practice it can be complicated to measure the impacts of a secondary event apart from its triggering event. A disaster classification that distinguishes primary and associated events can provide clear statistics on disaster impacts.		
Standardized human and economic impact definitions are the basis for consistent and compatible databases. In EM-DAT, the number of people killed is defined as 'Persons confirmed as dead and persons missing and presumed dead', and includes as such the missing people. The three DesInventar databases define fatalities as 'the number of people who died due to direct cause, whether immediately or sometime after the disaster'. The incorporation of a data element 'Number of people missing' into the EM-DAT database would facilitate interoperability between the databases.		
Characteristics of the disaster, such as the magnitude, event name, and the local time, can support the understanding of a disaster that has struck a population. EM-DAT provides these characteristics and could complement the information stored in the other databases.		
The physical impact of disasters (e.g., infrastructure, lifelines, buildings and housing, and the levels of damage) is currently difficult to capture in a global database. These impacts are more easily assessed and stored in the national disaster databases, such as DIMS. This information could complement data in the EM-DAT database.		
The duration of a disaster event is an important feature that determines the full impact of a disaster. In parallel with the duration (start date – end date) of a disaster, the duration of the disaster impact is questioned. For example, when do the homeless people have shelter? Measurement of the duration of the disaster and its impact should be clearly defined and consistently applied within the database. Providing information on start and end dates in the database facilitates comparability between datasets.		

Nepal (DIMS)	Indonesia (DIBI)	Sri Lanka (SDIS)
	The disaster area should ideally be represented by the disaster footprint, i.e., a polygon, or the coordinates of the affected cities. Although recent efforts have addressed the issues of attributing disaster footprints to disaster location, currently there is no standardized method used by international and national databases. DIBI has developed a system for linking its data to GIS software to provide maps of affected areas. EM-DAT can use a similar system of transferring its data to GIS software. Currently, this step needs human intervention to identify the latitude and longitude of the disaster areas. The storing of coordinates, or ideally disaster footprints, could benefit the compatibility between databases.	

4.3. ANALYSIS OF DATABASE COMPLETENESS

The completeness of information reported in the database is studied by calculating the proportion of total records containing information on human impact and on economic impact. The completeness of records is an indicator of quality because the reporting of events is only valuable for disaster risk management if full recording of disaster impact indicators is provided. Completeness of information in disaster loss databases depends to a large extent on the information provided by the data sources. The following tables provide a summary of the database completeness.

Table 7 : Philippines - Calamidat database

Database	Selected indicators	No. records	%
EM-DAT	**Total no. records**	**685**	**100%**
	no. empty (human impact)*	10	1.4%
	*no. empty (human + econ.**)*	7	1%
	*no. empty (human + econ. + physical***)*	n.a.	n.a.
Calamidat	**Total no. records**	**590**	**100%**
	no. empty (human impact‡)	84	14%
	no. empty (human + econ. ‡‡)	84	14%
	no. empty (human + econ. + physical‡‡‡)	n.a.	n.a.

Table 8: Bangladesh – DIDB database

Database	Selected indicators	No. records	%
EM-DAT	**Total no. records**	**477**	**100%**
	no. empty (human impact)*	16	3.3%
	*no. empty (human + econ.**)*	16	3.3%
	*no. empty (human + econ. + physical***)*	n.a.	n.a.
DIDB	**Total no. records**	**71**	**100%**
	no. empty (human impact‡)	44	62%
	no. empty (human + econ. ‡‡)	44	62%
	no. empty (human + econ. + physical‡‡‡)	63	89%

Table 9: Vietnam – DMC database

Database	Selected indicators	No. records	%
EM-DAT	**Total no. records**	**221**	**100%**
	no. empty (human impact)*	5	2.3%
	*no. empty (human + econ.**)*	3	1.4%
	*no. empty (human + econ. + physical***)*	n.a.	n.a.
DMC	**Total no. records**	**184**	**100%**
	no. empty (human impact‡)	2	1%
	no. empty (human + econ. ‡‡)	22	12%
	no. empty (human + econ. + physical‡‡)	24	13%

Table 10: Nepal – DIMS database

Database	Selected indicators	No. records	%
EM-DAT	**Total no. records**	**144**	**100%**
	no. empty (human impact)*	6	4.7%
	*no. empty (human + econ.**)*	5	3.5%
	*no. empty (human + econ. + physical***)*	n.a.	n.a.
DIMS	**Total no. records**	**16879**	**100%**
	no. empty (human impact‡)	6362	38%
	no. empty (human + econ ‡‡.)	4271	25%
	no. empty (human + econ. + physical‡‡)	2077	12%

Table 11: Indonesia – DIBI database

Database	Selected indicators	No. records	%
EM-DAT	**Total no. records**	**578**	**100%**
	no. empty (human impact)*	10	1.7%
	*no. empty (human + econ.**)*	7	1.2%
	*no. empty (human + econ. + physical***)*	n.a.	n.a.
Dibi	**Total no. records**	**6110**	**100%**
	no. empty (human impact‡)	3741	61%
	no. empty (human + econ. ‡‡)	3231	53%
	no. empty (human + econ. + physical‡‡)	782	13%

Table 12: Sri Lanka – SDIS database

Database	Selected indicators	No. records	%
EM-DAT	**Total no. records**	**96**	**100%**
	no. empty (human impact)*	4	4.2%
	*no. empty (human + econ.**)*	3	3.1%
	*no. empty (human + econ. + physical***)*	n.a.	n.a.
SDIS	**Total no. records**	100,844	**100%**
	no. empty (human impact‡)	7014	**6.9%**
	no. empty (human + econ. ‡‡)	7009	**6.9%**
	no. empty (human + econ. + physical‡‡‡)	6941	**6.8%**

* deaths, affected, injured, homeless.
** total damages, reconstruction damages, insured damages.
***Although the information is entered into the EM-DAT database, it is not available in a format that allow statistical analysis
‡ killed, injured, missing, affected communes, affected households, total no. affected, people needing aid.
‡‡ estimated damages
‡‡‡ houses and assets, education, health, agriculture/forestry, irrigation, transportation, fisheries, communication, industry, construction, others.

5. RECOMMENDATIONS

This section presents recommendations to strengthen, harmonize, and increase the interoperability of selected disaster databases in the Asian region. These guidance notes are based on discussions and interviews conducted with database managers and staff of database hosting institutions in the six countries. Recommendations aim at improving the visibility, accessibility, and applicability of disaster databases at the national level. Box 5.1 lists proposed methodological interventions, and Box 5.2 gives guidance related to accuracy and reliability. In Boxes 5.3 and 5.4, serviceability and accessibility are addressed, followed by sections on credibility (Box 5.5). This section ends with recommendations concerning the prerequisites and sustainability of the database (Box 5.6).

Box 5.1 Methodology

- ***Entry criteria:*** The use of entry criteria and thresholds related to the disaster impacts delineates when a disaster should be included in a database. It increases database homogeneity and avoids overrepresentation by disaster entries lacking human or economic impacts data, which may introduce biases when data are analysed.
- ***Disaster classification:*** The use of a hierarchical disaster type classification allows querying and sorting data on higher and lower scales and generating different levels of analysis.
- ***Definitions and standards:*** Use of internationally recognized definitions and standards increases data accuracy, facilitates data compilation, and allows inter-operability. In general, a clear listing and definition of disaster types contributes to data accuracy and reliability.
- ***Impact variables and database fields:*** Variables should be defined according to database objectives.
- ***Triggered events:*** Distinguishing primary and secondary 'triggered' disasters increases accuracy in the attribution of human impact data. It also allows for better insight into the complexity of the event.
- ***Identifier number:*** Use of a unique ID number—composed of, for example, by the year of occurrence, a sequential number, the type of disasters and the affected province, —is necessary for inter-operability of databases and (re-)aggregation of data. Through the use of a unique disaster number per event, disaster occurrence and impacts can be analysed at the national and local levels without introduction of biases from double-counting disaster events.
- ***Geocoding:*** A geocoding component integrated into the database system allows creation of digital maps and spatial analysis. It also allows for visualization of the impact of disasters in specific countries and/or

regions and contributes to providing enhanced products for policy and planning purposes.

- ***Data analysis**:* Procedures for data extraction and analysis must be developed and adapted according to specific internal and external use of the data.

Box 5.2 Accuracy and Reliability

- ***Accuracy and reliability of data sources:*** Cross-checking with additional sources of information (e.g., NGOs and the press, but also satellite-based images of the disaster impact) should be seen as an additional value for ensuring the accuracy of the data as well as the completeness of the information.
- ***Validation:*** Validating disaster data is key to maintaining a sound database. Implementing and strengthening the validation process ensures the quality and reliability of the data.
- ***Training and guidelines:*** Training staff and providing guidelines are important tools, especially in institutions where many persons are involved in the database management. It will facilitate data management and ensure data quality in different work phases, from information collection until output production, including data entry, validation, and quality-control processes.

Box 5.3 Serviceability

- ***Relevance and user profiles:*** Interaction with the end users of the database allows tailoring of database outputs and responding to user needs.
- ***Analytical capacities:*** The development and production of output products from the database ensures visibility, not only at a national but also at an international level.
- ***Visibility:*** User documentation, including guidelines and explanatory notes, should be clearly stated on the database's website to enhance data interpretation and reinforce the credibility, integrity, and professionalism of the database management board.
- ***Outputs and functions:*** Development of further online querying tools and outputs is an additional value for users.
- ***Timeliness:*** The applicability of a database could be strengthened by regular and timely data entry and dissemination, which will prevent missing or incomplete information and lead to increased data accuracy and completeness.

- ***Collaboration network:*** Reinforcing the collaboration network within the country or with international institutions that collect data allows completion and cross checking of information.
- ***Languages:*** Database websites should be developed in local languages as well as in English to increase applicability and horizontal and vertical interoperability.

Box 5.4 Accessibility

- ***Accessibility*****:** Facilitating the accessibility of the database website and extraction of the data increases database visibility and the use of valuable information.
- ***Analytical capacity***: Providing and developing analytical capacity increases the applicability of the database at the national and international levels.
- ***Information*****:** Contact details and any relevant information on the database must be easily accessible on the website.
- ***Institutional framework*****:** A non-hierarchical institutional structure facilitates exchanges and interoperability.

Box 5.5 Credibility

- ***Expertise and knowledge sharing*****:** It is beneficial for each institution to organize and/or attend disaster-databases–related conferences or workshops to exchange and share experiences and knowledge.
- ***Transparency*****:** Sharing information about database goals and objectives, methodology, concepts, definitions, data sources, and limitations reinforces credibility, integrity, and professionalism.
- ***Management and authority body*****:** The management body and/or authority should support quality improvement of the database over the long term.
- ***Impartiality*****:** The reporting of data by different information sources should ideally be done with transparency. When administrations are sensitive because of status or funding issues, the received data should be validated against other sources. Furthermore, efforts to desensitize administrations should be made to create trust in and compliance with the task of data and information sharing.

Box 5.6 Prerequisites and sustainability

- ***Resources:*** Adequate resources in terms of staff training, analytical capacities, database development and maintenance, additional programming to ensure outputs and services to the users, data collection and validation, and data analysis and reporting should be assured.
- ***Laying foundations:*** The minimum required technological infrastructure implemented at the national and local levels should be compatible with the database system.
- ***Continuity:*** The institutional framework should guarantee sustainability with budget allocation from the national government.
- ***Collaboration network:*** Joint efforts within the country as well as internationally to support country-level needs, share experiences, and exchange data and solutions, as well as building more analytical capacity, would benefit all countries. This collaboration network could be established between national disaster databases in the region and a global disaster database such EM-DAT.

6. GUIDELINES

One of the major current challenges in the field of disaster data is to overcome the limitations induced by the lack of clear standards and definitions, which often leads to reduced reliability and poor interoperability of different disaster data compilation initiatives. CRED has argued for years for the creation of internationally recognized standards and definitions. The present guideline is one step further toward this goal.

Based on the study conducted in six Asian countries and the conclusions and recommendations of the individual country assessment reports, this last section consists of a guideline on good practices for the development of disaster loss databases worldwide. This guideline covers all issues identified in the quality framework and the important steps in the development of a disaster database. The target group is database managers involved in daily database management and related practical work.

Common words recur in discussions about disaster databases: comparability and interoperability. Identifying disaster events across different databases at the global, national, or sub-national levels remains challenging, but this guideline can be seen as an important tool in accomplishing this goal.

The guideline is structured in chronological order, addressing data collection and sources, database structure, data entry, data validation, data analysis, database outputs, and dissemination. Each section provides key items to directly build upon when implementing or strengthening a disaster database.

This section provides general guidelines for the development of disaster databases and the compilation of reliable data, serving as practical guiding principles for the creation and maintenance of natural disaster impact databases worldwide.

6.1. DISASTER DATA COLLECTION AND DATA SOURCES

- It is important to identify relevant and reliable sources of information within the appropriate governmental agencies but also within the media (press) and international organizations, including NGOs located within the countries (e.g., UNDP, National Red Cross and Red Crescent Societies). It is also important to collect information from official and recognized data sources, but in addition, cross-checking with additional sources of information ensures data completeness and reliability. For example, it is useful to identify at least one scientific source of information (depending on the type of event) which provides more in depth information of the nature of the events (e.g. USGS for earthquakes)

- Primary sources or secondary sources can be used to collect data. Primary data sources are the agents that collect data by direct observation in the field (e.g., government agencies observing damage to infrastructure). Secondary sources are the institutions that gather data from the primary data sources and (dis)aggregate or summarize the information before making it public (e.g., an international relief organization distributing event reports to inform the humanitarian community).

- If the database is compiled from different sources of information, this fact must be specified to database users. A method of ranking the sources according to their reliability or completeness will help when conflicting information is provided. A quality indicator can be added to the source used, it will allow identifying automatically weak datasets. The use of multiple sources strengthens database reliability and provides complementary information, but it must be ensured that the information is provided at similar administrative levels.

- Data sources should provide acceptable coverage of information, in terms of

    - disaster types;

    - geographical area: the regions or parts of the country from which data are compiled in the database;

    - level of resolution: the level of resolution refers to the level of aggregation at which data are presented. Global observers like EM-DAT collect and present data as national level aggregates. National

observers collect and present data at provincial, municipal, or higher resolutions. Urban observers disaggregate data at the neighbourhood, block, or household levels;

- level of observation: the level of observation refers to the sources of information that agents use in collecting loss data. At the global level, observers rely on communications from international aid organizations or central government agencies. At the national or sub-national levels, usually local governments, field information, and local media reports are used; and
- time period.

- The frequency of data provision should be appropriate to the objectives of the database.
- The development and use of a standard data collection form for national and sub-national administrations, which includes commonly accepted and understood terms and definitions as well as a guideline, increase data accuracy. The use of a standard data collection form, preferably from the creation of the database onwards, ensures data consistency in the database.
- Explanatory notes should accompany the standard collection form and include clear definitions for the variables and about the information to be collected.
- Training of staff involved in the data collection ensures complete comprehension of the information that needs to be collected and contributes to the accuracy of the collected data.
- The updating of data has to be seen as an important step in the process of collecting data, as newly released information may become available a long time after an event occurs. In addition, long-term events, such as droughts, may need to be monitored for several years. The updating date associated with the source may be integrated into the database as an additional indicator to retrieve the information.
- The data sources and information reports should be archived, available for re-checking or addition of information to the database. This archive can be paper or electronic.

- The quality of disaster databases can only be as good as the reporting system that feeds them. Therefore, having an internationally recognized and accepted system for collecting data is an essential tool.

6.2. DATABASE STRUCTURE

- The structure of the database must be developed with a long-term vision before being implemented and be accompanied by complete documentation to allow for and facilitate future modifications or database development. One important step when building a database is to have a clear structure of the flow of the information and of the different types of information that need to be filled in at different levels in the database. It is important to understand that the basic structure of the database must reflect the different stages in which the information is being entered. A clear structure should be designed beforehand and applied from the beginning to avoid "retrofitting work", which is a time-consuming task.
- The architecture of the database must be based of the 3 following elements: (i) defining how the users understands the organization of the data; (ii) defining how the data is physically stored and processed and (iii), the level of inter-action between the two first elements.
- The variables used in the disaster database should fit: (i) the objective(s) of the database; (ii) the information needs of the managing institute or the country; and (iii) the information available from data sources that are used. The data fields that are included in a database are therefore a compromise of these conditions. It is useless to add additional indicators or fields that might not be completed because of missing information. Data fields that are not further used for analysis because of unavailable information or the lack of need for related outputs should be removed or adapted to make the database more clearly structured and easier to manipulate.
- The following essential fields are identified: Unique disaster ID number; Disaster type and sub-type; Location; Event Start and End dates; Human impact indicators (number of deaths, missing, injured, affected); and economic impact indicators (general and direct impacts are a minimum and

if possible breakdown by sectors can be given). The database structure should allow aggregation and disaggregation of information.

- Technical standards are needed to establish mechanisms for disaster-data–based computer systems, to interoperate between them; in other words, to specify the protocols by which computer systems will interact and speak a common language. Standards permit comparability of data from different sources to improve and facilitate analysis and allow integration of multiple sources of data on disasters. This integration may be horizontal (integration of data from different geographic areas or from different events, different times, different themes, or in general from different dimensions of the disaster data), or vertical (e.g., different levels of geographical resolution, from main disaster groups to specific types and vice versa).
- The system should be developed in a simple manner to allow easy management.
- The database should be adaptable over time. The structure should be flexible and also compatible with other existing systems.
- Backups of the database should be made regularly and stored in safe environments.

6.3. DATA ENTRY

6.3.1. IDENTIFICATION OF THE DISASTER EVENTS

- A clear definition of a "disaster event" (Hazard vs. Disaster) is crucial. Whether there are entry criteria or not, a definition is necessary to avoid entering a series of "disaster" events that have no human or economic impact. The definition of a disaster is much debated, and different agencies have issued various versions, largely according to their vision and objectives. For the purpose of a database, a working definition must be formulated to help the data manager and technical staff in deciding what would constitute a valid case of entry.
- The inclusion criteria for disaster events are usually linked to the database objectives. If criteria thresholds are fixed, they should be measurable in a quantitative and/or qualitative way.

Example: Quantitative threshold: 10 killed and/or 100 affected
Qualitative threshold: declaration of a state of emergency

- Attribution of a unique ID number: This identifier is a unique disaster number assigned for each event allowing identification of each individual record in the database. The use of an ID number allows the aggregation and disaggregation of disaster events. Based on the EM-DAT model, for example, an ID number system in a national and sub-national database could be as follows:

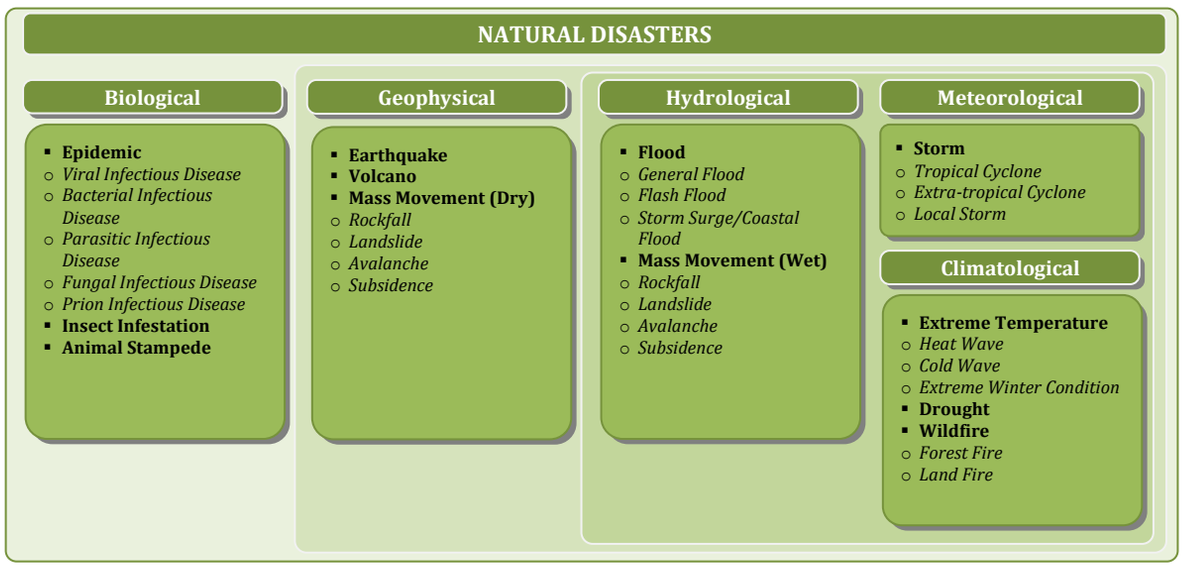
Example: Earthquake China, 30 August 2008—Miyi, Lihui, Panzhihua districts (Sichuan province); Kunming, Chuxiong, Yuanmo districts (Yunnan province)

Year	*SeqNumber*	*DisasterCode*	*Admin1LevelCode*	*Admin2LevelCode*
2008	*00374*	*EQ*	*SIC*	*Miy*
2008	*00374*	*EQ*	*SIC*	*Lih*
2008	*00374*	*EQ*	*SIC*	*Pan*
2008	*00374*	*EQ*	*YUN*	*Kun*
Etc…				

- Differences and a lack of standardization of typology and taxonomy complicate comparison of data sets. Therefore, disaster typologies should be clearly defined. In addition, databases struggle with disaster (sub) type classifications as well as their primary and secondary effects (or associated disasters). Without standardized terminology, databases continue to face decreased precision in reporting disaster-related impacts. Having a classification is useful for conducting analyses and aggregating or disaggregating disaster event data: i.e., a hierarchical classification (aggregation and disaggregation of disaster type, sub-type, group). CRED and MunichRe have recently (2009) published a working paper that proposes a standard disaster category classification and peril terminology[15]. Each database may have its own specificity, and a disaster event may be classified differently from one database to another according to the methodology (i.e., cyclone or flood, earthquake or tsunami) or the typology (winter storm or cold/frost). The goals of the initiative undertaken by CRED and MunichRe are to (i) create and agree on a common hierarchy and terminology for all disaster loss databases; and (ii) establish a common, agreed-upon definition of disaster groups, main types, and sub-types that is simple and

[15] Below R., Wirtz A., Guha- Sapir D. Classification and peril terminology for operational purposes - Common Accord between CRED and MunichRe, October 2009 (http://cred.be/sites/default/files/DisCatClass_264.pdf).

self-explanatory. The proposed classification is based on "triggering event" logic, within a hierarchical classification.



- Another important aspect is to link the triggering event and the associated disasters (e.g., flood/landslides; drought/forest fires); the information can be put into an additional field (e.g., Associated Disasters) to have a complete overview of a disaster event and its consequences, which may vary from one region to the other. For example, a storm may lead to flood in one part of a country and landslides in another area. This strategy will also avoid duplication of disaster events.

6.3.2. IDENTIFICATION OF GEOGRAPHICAL AND TEMPORAL INFORMATION

- Location: According to the level of database resolution, all geographical specifications (e.g., name of city, village, department, district, province, state) must be included in the database. This inclusion allows for the development of a subsequent analysis of disaster occurrence and impact by region, district, or any other sub-national administrative boundary. A methodology should be applied when entering location information, maintaining a standard procedure on entering this information, including all levels of resolution (e.g., working from the highest resolution to the lowest or the inverse).

- Temporal aspect: Defining the start and end dates of a disaster allows measurement of the duration of a disaster event. It is important to have a stringent definition of the start and the end date of a disaster and to use this consistently in the database for all disaster.
    - Start month/day/year: The date the disaster occurred, which is well defined for all sudden-impact disasters (e.g., earthquake), but for a disaster developing gradually over a longer time period and also geographically (e.g., drought), there is no specific onset date; in such cases, a specific methodology can be applied and the field "day" could be left blank.
    - End month/day/year: The date when the disaster ended; as for the start date, the end date is well defined for all sudden-impact disasters, and the same rule will apply for the long-term disaster events.

6.3.3. IDENTIFICATION OF THE HUMAN AND THE ECONOMIC IMPACT

- The definition of the human and economic impact has to be established. There is no absolute definition, but each database manager must provide the definitions to the database users. Commonly, the main/obligatory indicators used in a disaster loss database are as follows:
    - Human impact: deaths, missing, injured, homeless, affected
        - Secondary indicators could be those evacuated, victims, etc.
    - Economic impact: direct and indirect costs, separately recorded and not mixed in different data sets
        - Secondary indicators could be aid contribution, insured losses, reconstruction costs, as well as deflation/inflation tool.

The poor frequency of reported economic losses is notably the result of problems related to damage assessment. Although standard methods for assessing economic

losses exist (e.g., ECLAC – Economic Commission for Latin America and the Caribbean), there are no internationally accepted methods that can be used by any country and across all disaster types to measure both direct and indirect costs.

6.3.4. DATA VALIDATION

- Before making the data publicly available, it is important to set up a validation process and implement a quality-control system. This quality control aims at
    - avoiding or correcting typing errors and duplicate values;
    - completing the information if necessary;
    - checking extreme values; and
    - assessing missing data.

  Cross-checking with other sources of information (e.g., the press, NGOs) and also with other data source information (e.g., population) is an important step of the validation process. The validation process targets avoiding errors but also allows comparison of data coming from other sources of information.

- An internal cross-error checking routine will provide an alert if the information is not entered properly or if some data appear suspect (e.g., the number of affected people is higher than the total population of a specific province or district).

- A regular and timely update (for example, once every three months) of disaster data should be made to provide up-to-date information for internal and external use.

6.3.5. DATA ANALYSIS

- The compilation of disaster data in itself has no meaning if the data are not analysed or if no useful outputs are being produced. Therefore, the development of analytical tools is an integral part of database quality management. Performing analysis based on the compiled data in the database is important to generate and share new knowledge for use by policy makers and all stakeholders in disaster risk management. The quality and completeness of the data are critical for the usefulness of the database system. The analysis must contribute to presenting information that is relevant and can be used as an evidence base for policy and stakeholder decisions.

- Disaster data analysis supports:
    - better understanding of risk and vulnerability patterns;
    - measuring impacts on the population, economy, and environment;

- identifying characteristics of disasters and trends; and
- evaluation of disaster loss-reduction efforts.

- Data can be analysed in different ways according to the final users of the information:
  - Institutional (summarized information per region)
  - Research (detailed information, publications)
  - Media (timely information, major disasters, summaries, historical aspects, trends)
  - Consultancies

- Once the database is established and the data fully validated, the main questions before starting analysis are as follows:
  - How are the data used?
  - How can knowledge be generated and analytical capacities enhanced?
  - What kind of analytical methodologies can be used?

  Governments, hosting, or funding institutions put significant efforts into creating and maintaining disaster databases with the aim of providing useful and reliable outputs. Thus, asking the right questions will help in developing appropriate analytical tools.

- The organization of information depends on the purpose and target group and should combine at least the three main aspects:
  - Geographical: specific location, comparison of different locations, aggregation at the national level
  - Type of disaster: specific disaster type, comparison of different disaster types, general analysis
  - Type of impact: occurrence, human, and economic impact

- The main types of statistical treatments are as follows:
  - Descriptive statistics (e.g., absolute numbers, totals, averages)
    Example: Number of people affected by natural disasters in the past two decades
  - Trend analysis
    Example: Trends and forecasting of the human impact of natural disasters
  - Advanced treatment
    Example: Multivariate statistical analysis
  - Composite indicators (e.g., additional tables, Gross Domestic Product, population)
    Example: Number of victims per 100,000 inhabitants by income group

6.3.6. DATABASE OUTPUTS AND DISSEMINATION

- Once the database is established, the data fully validated, and analytical tools developed, the next step is to provide database users with easy access to data and database outputs. However, one should keep in mind that the compilation of a complete, valid, and accurate set of disaster data is more important than the development of analytical tools or a database website because these have no use if the data behind them are not reliable.

- It is necessary to have a clear understanding of user needs and profiles to identify further development of database functions and outputs and implement a data access policy.

- The visualization of the information based on the database is important. Users should be able to access
    - pre-formatted products (e.g., disaster profiles, country summary tables);
    - query options for tailored data selection (e.g., by time period, disaster type, and/or geographical area); and
    - detailed information or raw data.

- Confidentiality issues must be addressed if there are any and clearly stated on the database website and in database policy regulations.

- A database should be freely accessible. If access is limited for policy or financial reasons, this fact should be stated on the website of the institution and in related documents.

- The promotion of a database and dissemination of database outputs are important for
    - branding and building a credible reputation;
    - enhancing the visibility of the database (national and international); and
    - increasing the number of users.

- A database logo, acronym, and reference should be consistently applied to allow for identification of the database and to build (inter)national recognition.

- The number of database users or number of times the database is accessed are important performance indicators that can be used as a justification for long-term maintenance of the database (funding purposes) and to create awareness of the need for maintaining a national disaster-loss database in risk-prone areas.

- Promotion and dissemination strategies should be developed. Examples of dissemination channels are as follows:
    - the website of the organization or a dedicated database website:
        - by posting general information on the database and description of objectives, guidelines, entry criteria, methodologies, and conditions of use, and
        - through supporting documents for online query and visualization tools;
    - reports, newsletters, press releases;
    - publications in scientific journals;
    - collaborative networks and international organizations; and
    - interactive and social media.

- Training of personnel and attendance at conferences, meetings, and workshops related to disaster data increase professional expertise and institutional credibility. These avenues provide opportunities to exchange ideas and best practices with similar initiatives and allow the creation of expert networks.

7. CONCLUSIONS

The effectiveness of disaster preparedness and prevention depends on the evidence base on which the programme is anchored. Equally importantly, disaster data is central to studies that link disasters to health, social systems, poverty or even climate change. While there is a growing recognition of the need for accurate and comparable data on the impacts of disasters, there is still much room for improvement.

Since many years, CRED has persevered in its efforts to improve the quality of data on human impacts of disaster, engaging in many methodological initiatives with collaborators. Today, with new and cheaper technologies, information on the human impact of disasters should be systematically recorded and harmonized for comparisons across regions and also against time.

Most importantly, we have realized that to design a framework to improve data quality, we needed to get down in the weeds and examine the way in which countries actually operate. We did not want to work towards a solution without a sound grasp of the issues faced by our national colleagues.

We conclude from our study that there is an urgent need for **robust field methodologies** to estimate the number of dead, injured, and affected and guidelines for their use by national governments, international policy setting organizations and relief agencies. Ambiguity in figures that encompass an unspecified variety of groups or conditions, a common problem in disaster impact reporting, significantly reduces the usability of the data. We strongly advocate the development of standard methodologies that every agency can use to prevent ambiguous data from becoming a source of misguided policy and erroneous decision-making.

Another key area for improvement is greater standardization of **data compilation methods and definitions**. This goal can be achieved only by joint international efforts to develop these tools and make them available for national-level use.

Finally, one of the limitations of our study is that we have explored only six databases with a specific context and therefore our conclusions may not necessarily apply to other scenarios. However, we feel that the lessons learnt from this exercise and from the experience of EMDAT significantly bring forward the discussion on global data harmonization and inter-operability.

8. BIBLIOGRAPHY

Below R., Wirtz A., Guha- Sapir D. Classification and peril terminology for operational purposes - Common Accord between CRED and MunichRe, October 2009 (http://cred.be/sites/default/files/DisCatClass_264.pdf).

CRED, Workshop to Improve the Compilation of Reliable Data on Disaster Occurrence and Impact: Workshop Proceedings, Bangkok, 2-4 April 2006 (http://emdat.be/sites/default/files/BangkokWorkshopProceedings.pdf).

CRED. Proposed principles and guidelines for the collection and dissemination of disaster related data, Report on the IERRIS Workshop, 7-9 September 1992.

FAO Statistical Data Quality Framework: A multi-layered approach to monitoring and assessment. Conference on Data Quality for International Organizations, Wiesbaden, Germany, 2004.

Hoyle D. ISO 9000 Quality Systems Handbook. Butterworth-Heinemann, 2006.

International Association for Information and Data Quality (IAIDQ) (www.iaidq.org/main/glossary.shtml).

International Monetary Fund (IMF) Data Quality Assessment Framework 2003 (www.dsbb.imf.org/Applications/web/dqrs/dqrsdqaf/).

Sapir D., Below R, Hoyois Ph. Data on disasters: Easier said than done, Disaster and Development, 2006; 1 (1): 155-63.

Sapir D.G. and Misson C. The Development of a Database on Disasters. Disasters, 1992; 16(1): 80-86.

Tschoegl L., Guha-Sapir D., Below R. An analytical review of selected data sets on natural disasters and impacts. Paper prepared for the UNDP/CRED Workshop on Improving Compilation of Reliable Data on Disaster Occurrence and Impact, Bangkok, 2–4 April 2006 (http://emdat.be/sites/default/files/TschoeglDataSetsReview.pdf).

United Nations Development Programme. Guidelines and lessons for establishing and institutionalizing disaster loss databases, UNDP, 2009.

World Bank Development Data Group and UNESCO Institute for Statistics. A framework for assessing the quality of education statistics,(www.uis.unesco.org/TEMPLATE/pdf/SCB/DQAF%20for%20education%20statistics.pdf).

ANNEX 1 OVERVIEW OF DATA ELEMENTS IN DATABASES (COMPARING TO EM-DAT) – STAND ALONE DATABASES

Table 1: Philippines - Calamidat database

Database	Data elements	
Present in both	**Indispensible**	**Optional**
	Disaster type	Insertion date
	Start year, month, day	GLIDE
	End year, month, day	N° Houses damaged
	N° deaths	N° Houses destroyed
	N° injured	Impact on infrastructure
	N° (total) affected	Local time
	Data sources	Aid contribution
	Geographic location	Disaster Magnitude/scale
	ID number	
	Name	
Not present in Calamidat (but present in EM-DAT)	**Indispensible**	**Optional**
	Associated Disasters	**Origin/causes**
	Disaster sub-type	Entry criteria
	Economic damages (US$)[16]	River basin
		N° homeless
		Disaster Group
		Latitude/Longitude
Not present in EM-DAT (but present in Calamidat)	**Indispensible**	**Optional**
	N° Missing[17]	Summary
	Losses in local currency[18]	Comments
		Last update
		Affected barangays
		Affected families
		Evacuation centers

[16] Only relevant for global databases
[17] Included in Deaths
[18] Only relevant for national and sub-national databases

		Families served inside
		Persons served inside
		Families served outside
		Persons served outside
		Assistance
		Asset deployments

Table 2: Bangladesh – DIDB database

Database	Data elements	
Present in both	**Indispensible**	**Optional**
	Event ID Number	GLIDE
	Date of event/Start date	Time of event
	Location	Houses damaged
	Deaths	Damaged crops (full)
	Affected people	Damaged crops (partial)
	Injured people	Damaged households (full)
	Disaster type	Damaged households (partial)
	Associated Disasters	Affected educational institute (full)
		Affected educational institute (partial)
		Damaged road (full)
		Damaged road (partial
		Damaged bridge/culvert
		Max. wind speed
		ID
Not present in DIDB (but present in EM-DAT)	**Indispensible**	**Optional**
	Disaster sub-type	Entry criteria
	End date	Aid contribution
	Source name	Disaster Group
	No. Homeless	Disaster Magnitude/scale
	Comment	River Basin
	Economic damages (US$)	Disaster sub-group

		Entered by
		Origin
		Appeal for international assistance
		Declaration of disasters
		Reconstruction cost
		Insured losses
		Disaster impact on sectors
		Latitude/Longitude
Not present in EM-DAT (but present in DIDB)	**Indispensible**	**Optional**
	Missing	Duration
	Losses in local currency	Remarks
		Geographic coverage[19]
		Affected families
		N° affected
		Affected population
		Affected Districts
		Trees destroyed
		N° livestock death
		Damaged Embakment
		People took shelter
		Lost/damage
		Nature of the phenomenon
		Geom
		Damage info

Table 3: Vietnam – DMC Database

Database	Data elements	
Present in both	**Indispensible**	**Optional**
	Disaster type	Insertion date
	Start year, month, day	No. homeless
	No. deaths	
	No. injured	
	No. (total) affected	
	Data sources	
	Economic damages	

[19] Geographic coverage = coding

	Geographic location	
Not present in DMC (but present in EM-DAT)	**Indispensible**	**Optional**
	Duration or end year/ month/ day	Notes
	Disaster identification code	Disaster group
	Associated disasters	Origin/cause
	Economic damages (US$)	Common name of event
		Entry criteria/reason for entering
		Local time
		Aid contribution
		Disaster magnitude
		River basin
		Latitude/longitude
Not present in EM-DAT (but present in DMC)	**Indispensible**	**Optional**
	No. missing	Province
	Losses in local currency	No. communes affected
		No. households affected
		No. children deceased
		No. people in need of relief
		No. houses and assets (houses, offices, ..)
		No. education (schools, class rooms, ...)
		No. health. (hospitals, health centres, ...)
		No. agriculture/forestry (rice fields, seeds..
		No. irrigation (dykes, canals,...)
		No. transportation (roads, bridges, ...)
		No. fisheries (Fish areas, fish boats,...)

		No. communication (telephone wires, ..)
		No. industry (electric towers, engines,...)
		No. construction (sites, materials,...)
		Others (school desks, fertilizers,...)

ANNEX 2 OVERVIEW OF DATA ELEMENTS IN DATABASES (COMPARING TO EM-DAT) – DESINVENTAR DATABASES

Table 1: Nepal – DIMS Database

Database	Data elements	
Present in both	**Indispensible**	**Optional**
	Disaster type	Notes
	Start year, month, day	Serial number
	Duration or end year/ month/ day	Origin/cause
	No. deaths	Disaster magnitude
	No. injured	Insertion date
	No. (total) affected	Inserted by
	No. victims	
	Data sources	
	Economic damages	
	Geographic location	
Not present in DIMS (but present in EM-DAT)	**Indispensible**	**Optional**
	Disaster identification code	Disaster group
	Associated disasters	No. homeless
	Economic damages (US$)	Common name of event
		Entry criteria/reason for entering
		Local time
		Aid contribution
		River basin

		Latitude/longitude
Not present in EM-DAT (but present in DIMS)	**Indispensible**	**Optional**
	No. missing	Region
	Losses in local currency	District
		Village
		Region GIS code
		District GIS code
		Village GIS code
		No. destroyed houses
		No. affected houses
		Affected routes (m)
		Farming and Forest (Ha)
		Livestock
		Education Centers
		Medical Centers
		Infrastructure*
		Other losses
		No. relocated
		No. evacuated

*Infrastructure includes non-numerical damages to transport, agriculture, communications, power, education, relief, water supply, sewerage and drainage, industry, health, others.

Table 2: Indonesia – DIBI Database

Database	Data elements	
Present in both	**Indispensible**	**Optional**
	Disaster type	Notes
	Start year, month, day	Serial number
	No. deaths	Aid contribution
	No. injured	
	No. (total) affected	
	Data sources	
	Economic damages	
	Geographic location	
	Associated disasters	
Not present in Dibi (but present in EM-DAT)	**Indispensible**	**Optional**
	Disaster identification	Disaster group

	code	
	Duration or end year/ month/ day	No. homeless
	Economic damages (US$)	Common name of event
		Entry criteria/reason for entering
		Origin
		Disaster magnitude
		Local time
		River basin
		Latitude/longitude
Not present in EM-DAT (but present in Dibi)	**Indispensible**	**Optional**
	No. missing	Province
	Losses in local currency	District
		Province GIS code
		District GIS code
		Latitude/longitude
		No. damaged houses (heavily/lightly)
		No. health facilities
		No. education facilities
		Rice fields
		Road
		No. evacuated
		No. inundated houses
		No. worship facilities
		No. offices
		No. kiosks
		Infrastructure
		No. plantations
		No. ponds
		No. irrigation facilities
		No. buildings

Table 3: Sri Lanka – SDIS Database

Database	Data elements	
Present in both	**Indispensible**	**Optional**

	Disaster type	Aid contribution
	Start year, month, day	Houses damaged
	No. deaths	Houses destroyed
	No. injured	Damaged in roads
	No. (total) affected	Damaged in crops
	Data sources	Lost cattle
	Economic damages	Education centers
	Geographic location	Hospitals
	Associated disasters	Affected sectors
	Sources	
	Duration or end year/ month/ day	
	Magnitude	
Not present in SDIS (but present in EM-DAT)	**Indispensible**	**Optional**
	Disaster identification code	Disaster group
	Disaster sub-type	No. homeless
	Economic damages (US$)	Common name of event
		Entry criteria/reason for entering
		Origin
		Disaster magnitude
		Local time
		River basin
		Latitude/Longitude
Not present in EM-DAT (but present in SDIS)	**Indispensible**	**Optional**
	N° Missing	Serial number
	Losses in local currency	N°Evacuated
		N°Relocated
		N°Victims
		Infrastructure
		No. plantations
		No. ponds
		No. irrigation facilities
		No. buildings

ANNEX 3 PREPARATORY QUESTIONNAIRE AND INTERVIEW

PART ONE:

General information on disaster database

Please fill in the following questions:

A. Database denomination[20]
A.1 Official name of database:

A.2 Acronym of database:

A.3 Description of database:

A.4 Objectives and purposes of the database:

A.5 For who is the database developed:

B. Information on the institute*
B.1 Institute in charge of hosting the database:

B.2 Institute complete post address:

B.3 Institute URL address:

B.4 Institute Contact person/ Director:

C. Database contact information*
C.1 Database contact person:

C.2 Email:

C.3 Telephone number:

C.4 Unit in charge of the management of the database:

C.5 Complete address of unit (if different from the hosting institute):

[20]* More information in Part 3: Explanatory Notes

C.6.1 Website available to consult database: ☐ Yes ☐ No

C.6.2 If yes, what is the database URL address (if different from the hosting institute):

C.7 Languages of the offline (internal working) database:

C.8. Languages of the online database:

D. Database content description*

D.1 Level of observation*:

D.2 Level of resolution*:

D.3 Geographical coverage*:

D.4 Types of disasters monitored:

Natural disasters: ☐ Yes ☐ No
Technological disasters/Man-made disasters ☐ Yes ☐ No
Complex emergencies: ☐ Yes ☐ No
Other:
.....
.....

D.5 Period covered (start year-end year):

D.6 Still active: ☐ Yes ☐ No, end date:
☐ No, but will start:

D.7 How often are data updated in the database:

D.8 Latest update of data in the database (dd/mm/yyyy):

D.9.1 Identification code present for each disaster: ☐ Yes ☐ No

D.9.2 If yes, is it unique code or are multiple ID numbers assigned to one event?

D.9.3 how is the ID defined?

D.10.1 Are data aggregated of disaggregated? ☐ Yes, aggregated
☐ Yes, disaggregated
☐ No

D.10.2 If yes, at which level are data (dis)aggregated?

D.11 Total number of records present in offline database:

D.12 Total Number of records present in online database:

D.13 If website available to consult database: Is this an interactive web-based system*: ☐ Yes ☐ No

D.14 Content of website*:

- static tables/ list of events ☐ Yes ☐ No
- dynamic querying ☐ Yes ☐ No
- creating charts ☐ Yes ☐ No
- statistical tools ☐ Yes ☐ No
- creating maps ☐ Yes ☐ No
- exportable data ☐ Yes ☐ No
- access to full event reports ☐ Yes ☐ No
- access to original report of data source ☐ Yes ☐ No
- others:

.....

D.15 Additional links or portals through which the database is available (besides database URL):

D.16 List of outputs produced (e.g. publications, statistical reports):

D.17 General definition used for 'disaster event':

D.18. Is there any disaster classification/hierarchy implemented in your database?* ☐ Yes ☐ No

D.19.1 Disaster groups present*: ☐ Yes ☐ No

D.19.2 If yes, what is the definition of each disaster group?

D.20.1 Disaster subgroups present*: ☐ Yes ☐ No

D.20.2 If yes, what is the definition of each disaster subgroup?

D.21.1 Disaster types present*: ☐ Yes ☐ No

D.21.2 If yes, what is the definition of each disaster type?

D.22.1 Disaster subtypes present*: ☐ Yes ☐ No

D.22.2 If yes, what is the definition of each disaster subtype?

D.23.1 Are your definitions based on international recognized standards?
☐ Yes ☐ No

D.23.2 If yes, which one(s)?

D.24.1 Do you have definitions of human impact indicators*:
☐ Yes ☐ No

D.24.2 If yes, please give the definitions for the following indicators:

- 'affected people':
- 'fatalities' (deaths/people killed):
- 'injured people':
- 'missing people':
- 'evacuated people':
- 'homeless people':
- 'victims':
- 'relocated people':
- *Please note down other indicators not included above, as well as their definitions:*

  .....
  .....

D.25.1 Do you have definitions of economic or structural impact indicators*:
☐ Yes ☐ No

D.25.2 If yes, please give the definitions for the following indicators:

- 'economic damage':
- 'sector damage':
- 'infrastructure damage':
- ´aid contribution´:
- *Please note down definitions of indicators other than the ones described above:*

  .....
  .....

PART TWO:

Interviews

Name of person interviewed:
Function:

A. Methodology

Concepts and definitions

A.1.1 Do you have criteria for the inclusion of events in the database? ☐Yes ☐No

A.1.2 If yes, which one(s)?

A.2.1 Are your entry criteria based on (inter)national recognized standards? ☐ Yes ☐ No

A.2.2 If yes, which one(s)?

A.3.1 Do you have a disaster classification* ☐ Yes ☐ No

A.3.2 If yes, how are the disasters classified? (e.g. hierarchical or equal level)

A.4 Is your classification based on (inter)national recognized standards? ☐ Yes ☐ No

A.5.1 Do you distinguish events that are triggered by primary events? ☐ Yes ☐ No

A.5.2 If yes, how do you enter them into the database?

*Collecting data**

A.6 Do you have a standardized way of collecting information? (e.g. first general info, then detailed)

A.7 Can you give an example of how you collect the data in practice?

A.8 How many sources are providing data for the database?

A.9 Are you using primary sources or secondary sources or both*?

A.10 What are the major data sources for the database?

A.11 Can you give a list of the main sources you use?

A.12 In what format do data sources generally provide data?

Paper format ☐ Yes ☐ No
Electronic format ☐ Yes ☐ No
Other:
……
……

A.13 What is the frequency of data provision (for the main data sources)?

A.14 Do you use a limited defined set of data sources or do you also search for complementary sources if needed?

A.15 Are data from sources available to the public or is there limited access?

A.16 Are there financial costs for obtaining data from the data sources?

Entering data

A.17 How many people enter data into the database?

A.18 How often are data entered in the database?

A.19 What type of database do you have?

A.20 Is this a flat database (excel) or a relational database (access, MySQL)?

A.21 What types of information are inserted into the database:

- ☐ Information on the disaster (disaster type, name)
- ☐ Temporal information (date, year)
- ☐ Geographical information (location)
- ☐ Human impact information (# people killed)
- ☐ Economic impact information (aid contribution, direct economic losses)
- ☐ Infrastructural impact information (# houses destroyed)
- ☐ Causes or triggering events for the happening of disasters
- ☐ Others:

  ……
  ……

A.22.1 Do you insert georeferencing information for disasters into the database*? ☐ Yes ☐ No

A.22.2 If yes, which type(s) (points, lines, polygons)?

A.23 How many events are approximately inserted per year?

A.24.1 Do you archive the data source records after entering into the database? ☐ Yes ☐ No

A.24.2 If yes, how:

- ☐ Paper copies
- ☐ Electronic format

Technical orientation

A.25 How often are backups of the database made (IT)?

A.26 Where is system information being stored (IT)?

Analyzing data for internal use

A.27.1 Do you analyze data inserted in the database? ☐ Yes ☐ No

A.27.2 If yes, how do you extract the data from the database? (predefined queries, SQL, pivot table)

A.27.3 If yes, how do you analyze the data?

A.28 Which information fields from the database are used for further processing of the data (e.g. data table)?

A.29 Which software do you use?

A.30 For which goals and for whom are the analyzed data further used? (Distinguish internal/external)

B. Accuracy and reliability

Accuracy and reliability of data sources

B.1.1 Do your data sources provide data that are complete/detailed enough, or is information missing?

B.1.2 If missing: What information is missing?

B.2 What geographical area do the data sources cover?

B.3 Do you have difficulties understanding the terminology used by your sources?

B.4.1 Do you think that certain data sources can be biased? ☐ Yes ☐ No

B.4.2 If yes, how?

B.5.1 Do you check if data sources provide valid data? ☐ Yes ☐ No

B.5.2 If yes, how?

Accuracy and reliability of data in database

B.6 Do you have a standard data entry form*? ☐ Yes ☐ No

B.7 How many sources are used in order to enter one disaster event into the database?

B.8 How do you compare different sources for entering an event?

B.9 How do you manage conflicting information?

B.10.1 Do you rank the sources based on their quality? (e.g. priority sources) ☐ Yes ☐ No

B.10.2 If yes, in which way?

B.11 How do you treat missing values (empty data fields)?

B.12.1 Do you have a validation process? ☐ Yes ☐ No

B.12.2 If yes, how?

B.13.1 Are inserted data checked for duplicates: ☐ Yes ☐ No

B.13.2 If yes, how?

B.14.1 Are inserted data checked for typing errors? ☐ Yes ☐ No

B.14.2 If yes, how?

B.15 Does the database automatically flag strange or suspect data? (error message) ☐ Yes ☐ No

B.16 How often are data updated in the offline database?

B.17 How often are data updated in the online database (if applicable)?

B.18 Have you had the same data entry procedures for the database its start? (consistency) ☐ Yes ☐ No

B.19 Are there training possibilities to develop capacities of database staff? ☐ Yes ☐ No

B.20 Are there guidelines for data handling for internal (staff) use? ☐ Yes ☐ No

B.21 Which of the following fields must be completed in the database (minimum required fields):

- ☐ Date
- ☐ Disaster type
- ☐ Disaster ID
- ☐ Geographical information
- ☐ Human impact
- ☐ Economic costs
- ☐ Data source

Others:

……

……

B.22 Are these fields necessary for the system to introduce a new event? ☐ Yes ☐ No

B.23 Are these fields required for further analysis of the data?

B.24 What information fields could be useful to include in the database for use of further analysis?

C. Serviceability

Outputs and functions

C.1 Through what media do you disseminate the data (website, press,…)?

C.2 What outputs are made available for the user (reports, cd,…)?

C.3.1 Is the user able to aggregate/disaggregate data? ☐ Yes ☐ No

C.3.2 If yes, how?

Interpretability

C.4.1 Are charts and tables provided along with the data? ☐ Yes ☐ No

C.4.2 If yes, is information available to explain the charts/tables to facilitate their interpretation? ☐ Yes ☐ No

Timeliness/periodicity

C.5 At what intervals do you make data available to the public?

C.6 How much time is there approximately between an event occurring and the publication of data on the event?

Interoperability

C.7.1 Do you participate in a database network with other disaster databases? ☐ Yes ☐ No

C.7.2 If yes, Do you share a common objective? ☐ Yes ☐ No

C.7.2.1 If yes, what is this common objective?

C.7.2.2 With which databases do you share the objectives?

C.7.2.3 How regular are you in contact with the involved database managers/staff?

C.8.1 Have you ever exchanged data with other databases? ☐ Yes ☐ No

C.8.2 If yes, which one(s)?

C.9.1 Can your system export data? ☐ Yes ☐ No

C.9.2 If yes, in what format do you export data?

C.10.1 Can data be imported? ☐ Yes ☐ No

C.10.2 If yes, in what format do you import data?

C.11.1 Do you share standard data formats with other databases? ☐ Yes ☐ No

C.11.2 If yes, which formats?

C.12.1 Do you insert a GLIDE number for each event*? ☐ Yes ☐ No

C.12.2 If yes, since what year?

C.13.1 Do you share common classifications or terminology with other databases? ☐ Yes ☐ No

C.13.2 If yes, Which ones?

C.13.3 With whom?

C.14 Is data sharing (export) in any way restricted by your institution?

C.15.1 Would it be useful for you to exchange more information with other databases? ☐ Yes ☐ No

C.15.2 If yes, what information would be useful to share?

User documentation

C.16.1 Are guidelines or explanatory notes for the database available for the user? ☐ Yes ☐ No

C.16.2 If yes, in which format?

C.16.3 If yes, in what language?

C.17.1 Is there a Frequently Asked Question list available? ☐ Yes ☐ No

C.17.2 If yes, in which format?

C.17.3 In what language?

C.18.1 Is there access to publications and reports concerning the DB through the website? ☐ Yes ☐ No

C.18.2 If yes, in which format?

C.18.3 If yes, in what language?

Relevance and user profile

C.19.1 Are you in contact with a user advisory group? ☐ Yes ☐ No

C.19.2 If yes, how?

C.20.1 Do you consult an expert group for improving the database? ☐ Yes ☐ No

C.20.2 If yes, how?

C.21.1 Do you monitor user profiles? ☐ Yes ☐ No

C.21.2 If yes, have you reported the results? ☐ Yes ☐ No

C.21.3 If no, do you consider to conduct a user survey? ☐ Yes ☐ No

C.22.1 Do you monitor the satisfaction of the users? ☐ Yes ☐ No

C.22.2 If yes, how?

C.23 What information would you expect from a user survey?

C.24.1 Do you receive comments or suggestions from users? ☐ Yes ☐ No

C.24.2 If yes, what are the most common remarks?

C.25 What could be improved for the users of the database?

C.26 What is hindering to achieve this?

D. Accessibility

D.1 Are there costs to access the database? ☐ Yes ☐ No

D.2.1 Are there parts of the database that have restricted access? ☐ Yes ☐ No

D.2.2 If yes, Do users have to register? ☐ Yes ☐ No

D.2.3 Which parts of the data/database are openly accessible?

D.2.4 Which parts have restricted access?

D.2.5 What is the policy / criteria for giving or restricting access?

D.2.6 Is the access policy clearly stated on the website or available by other means to the users? ☐ Yes ☐ No

D.3 Are non-published data made available upon request? ☐ Yes ☐ No

D.4 Are contact details of the institute /database manager/ responsible available for users? ☐ Yes ☐ No

E. Credibility

Transparency

E.1 Do you have information on your website about:

- the institute ☐ Yes ☐ No
- goals and objectives of the database ☐ Yes ☐ No
- the database methodology ☐ Yes ☐ No
- definitions ☐ Yes ☐ No
- the data sources used for the database ☐ Yes ☐ No
- data analysis procedures applied ☐ Yes ☐ No
- data dissemination ☐ Yes ☐ No
- the limitations of the database ☐ Yes ☐ No
- recent changes in policies and practices concerning the database ☐ Yes ☐ No

E.2 Are data products identified by

- a logo ☐ Yes ☐ No
- an institute reference ☐ Yes ☐ No
- a citations ☐ Yes ☐ No

Expertise

E.3.1 Are publications/reports based on the database written by the institute? ☐ Yes ☐ No

E.3.2 If yes, which?

E.4.1 Are external publications/reports peer-reviewed? ☐ Yes ☐ No

E.4.2 If yes, which?

E.5.1 Are database-related conferences organized or attended? ☐ Yes ☐ No

E.5.2 If yes, which? (give some recent examples)

E.6.1 Are data published or cited in the media? ☐ Yes ☐ No

E.6.2 If yes, on any specific occasion(s)?

Quality management

E.7 Do you feel that the management supports the improvement of data quality? ☐ Yes ☐ No

Impartiality

E.8 Is your database used for resources allocations ☐ Yes ☐ No

F. Prerequisites and sustainability

*Institutional framework**

F.1 What is the institutional framework of the organization?

- ☐ Private
- ☐ Academic/research centre
- ☐ UN/International agencies
- ☐ Government
- ☐ NGO
- ☐ Other:

F.2 Is the database seen as the reference database in the country? ☐ Yes ☐ No

F.3 Could the database be maintained independently from its institutional framework? ☐ Yes ☐ No

Resources

F.4.1 Who are your: main funders?

F.4.2 secondary funders?

F.5 Is your funding sustainable on the long term? ☐ Yes ☐ No

F.6 How many staff currently work on the database?

F.7 What positions are lacking at the moment?

F.8 Which of the following costs are covered by the resources you have:

- ☐ data collection/insertion
- ☐ analysis of data
- ☐ computing systems/software
- ☐ office space and other fixed costs
- ☐ development of database products
- ☐ diffusion of data
- ☐ production of documentation and training materials

Collaboration network

F.9.1 Has your database been developed in collaboration with other institutions? ☐ Yes ☐ No

F.9.2 If yes, which one(s)?

F.10.1 Have you collaborated with other institutions to strengthen your database? ☐ Yes ☐ No

F.10.2 If yes, which one(s)?

F.11 With which institutions do you currently collaborate?

F.12 What is the goal of these collaborations?

F.13 Which collaborations would you like to develop in the future that could serve your database?

F.14 What do you need to develop these collaborations?

Continuity

F.15.1 What are the long term objectives of the database?

F.15.2 If any, are these supported by the hosting institute? ☐ Yes ☐ No

F.15.3 Are these supported by the funders? ☐ Yes ☐ No

F.15.4 Are these supported by your partnership network? ☐ Yes ☐ No

G. Additional remarks:

PART THREE:

Explanatory Notes

Database denomination

The database denomination is the name of the disaster database. The official name is the full name officially assigned to the database. The acronym is the short, commonly used name assigned to the database.

The description of the database provides some contextual information on the database, such as its development and daily management.

Information on the institute

The institute name is the complete name of the organization(s) that hosts the database.

Database contact information

This section includes the contact details of the unit and responsible person(s) that is in charge of the daily management of the database.

Database content description

- Level of observation
  - Level of observation refers to the sources of information that are used by the agents collecting loss data. At global level, observers rely on communications from international aid organizations or central government agencies. At national or sub-national levels usually local governments, field information and local media reports are used.

- Level of resolution
  - Level of resolution refers to the level of aggregation at which data is presented. Global observers like EM-DAT collect and present data as national level aggregates. National observers collect and present data at provincial, municipal or higher resolutions. Urban observers disaggregate data at neighbourhood, block or household levels.

- Geographical coverage
  - Geographical coverage refers to the regions or parts of the country on which data are compiled in the database

- Identification code
  - The identification number or code for each event record in the database. This can be a unique identifier assigned by the database

software, or additional identifiers developed by the staff for identification of the specific events.

- Interactive web-based system
  - Interactive web-based system refers to a website that can be interrogated by querying the data for specific and personalized information types.

- Content of website
  - Static tables: lists or summarizes events; no possibilities for further interrogation
  - Dynamic querying: users can interrogate the database by specific and personalized queries (questions)
  - Creating charts: users can create charts after defining their indicators
  - Statistical tools: users can develop their own study questions and perform analyses on selected data
  - Creating maps: users can create geographical maps containing disaster impact data
  - Exportable data: data or summarizing tables can be exported and further analyzed by the user
  - Access to full event reports: links are present that provide access to specific detailed information
  - Access to original data source: links are present that provide access to the original data from the sources

Disaster groups
In EM-DAT:
- Natural disasters
- Technological disasters

Disaster subgroups
In EM-DAT:
-Natural: Biological disasters
Geophysical disasters
Hydrological disasters
Meteorological disasters
Climatological disasters

Disaster types and subtypes
In EM-DAT:
Description of the disaster according to a pre-defined classification (for example, type: flood; sub-type: flash flood).
See: Annual Disaster Statistical Review, the numbers and trends, CRED, 2008 (www.emdat.be)

Disaster classification
The disaster classification is the way in which the different disaster events are ordered and defined.

Hierarchical: classification from broad categories to tailored disaster types
Equal: classification takes each event into account as an equivalent order

Example from EM-DAT: hierarchical classification
Disaster group: Natural
Disaster subgroup: Hydrological
Disaster type: Flood
Disaster subtype: Flash flood

Primary sources or secondary sources
Primary data sources are the agents that collect data by direct observation in the field (e.g. government agencies observing damage to infrastructure).
Secondary sources are the institutions that gather data from the primary data sources and (dis)aggregate or summarize the information before making it public (e.g. international relief organization distributing event reports to inform the humanitarian community).

Technical information
A backup is a copy of the database at a given moment, placed separately from the original database in order to secure retrieval of the product if the local system fails.
System information is all information on the structure and development of the database, including types of software used, data formats, scripts, and tables/fields present.

Standard data entry form
A standard data entry form is a pre-defined format for entering data into the database, covering the several types of information to be entered. This standard entry form prevents against errors during the data entry. Minimum required fields can be obliged to fill before the system can store the inserted event.

EM-DAT human impact indicators definitions

Number of people killed: Persons confirmed as dead and persons missing and presumed dead (official figures when available).

Number of people affected: People requiring immediate assistance during a period of emergency; it can also include displaced or evacuated people.

Number of people injured: People suffering from physical injuries, trauma or an illness requiring medical treatment as a direct result of a disaster.

Number of homeless: People needing immediate assistance for shelter.

Victims: Sum of killed and total affected.

Total number of people affected: Sum of injured, homeless, and affected.

EM-DAT economic or structural impact indicators definitions

Total economic damages: (in 1000 US$): Several institutions have developed methodologies to quantify these losses in their specific domain. However, there is no standard procedure to determine a global figure for economic impact. Estimated damage are given (000') US$.

Georeferencing: The process of referencing a map image to a geographic location, by providing geographic coordinates that represent a textual location description. In other words, numbers (latitude and longitude) are assigned to descriptions. The purpose is to allow for easy mapping and spatial analysis of the phenomena observed at these locations.

GLIDE number: the GLobal IDEntifier number is a globally common Unique ID code for disasters. The components of a GLIDE number consist of two letters to identify the disaster type (e.g. EQ - earthquake); the year of the disaster; a six-digit, sequential disaster number; and the three-letter ISO code for country of occurrence. So, for example, the GLIDE number for West-India Earthquake in India is: EQ-2001-000033-IND. (*source: www.glidenumber.net)*

Institutional framework

The institutional framework is the structure and environment in which the institution that maintains and manages the database is embedded.

September, 2010

Centre for Research on the Epidemiology of Disasters
CRED

USAID
FROM THE AMERICAN PEOPLE