Journal of Computer Science and Information Technology
June2017, Vol. 5, No. 1, pp. 15-25
ISSN: 2334-2366 (Print), 2334-2374 (Online)

Published by American Research Institute for Policy Development
DOI: 10.15640/jcsit.v5n1a2
URL: https://doi.org/10.15640/jcsit.v5n1a2

Integrated Risk Components in Data Modeling for Risk Databases

Washington Okori[1] & Joseph Obua[2]

Abstract

This study examined the suitability of elements such as terms, models and databases necessary in the development of suitable risk models. The elements are applicable at all stages of disaster intervention as preparedness, mitigation, response and recovery. For each of the stages cross-sectoral integration of information is applied in understanding the short and long term impact of a particular disaster. Integration enhances the ability of disaster managers to appropriately analyze risks, develop baseline risk profile and make better informed decision before and after disasters. The need to build risk databases based on appropriate data modeling procedure that includes conceptual data models, logical data models and the required data needed is corroborated in this study. Data model focuses on "real-world" representation of data as seen by users. Entity-Relationship methodology was applied since it takes less time to design, provides better quality of design and the design is easy to understand. The entities found to be of interest in the disaster risk model are hazard, exposure, vulnerability and loss. It is recommended that there should be integration within each of the elements such as terms, models and databases to ensure that countries are well equipped with high quality data and information to anticipate and avert disasters such as flood. The data model should be well structured with the aim of improving performance of the system to enhance timely information retrieval.

Keywords: Data Model, Disaster, Integration, Risk Database, Risk Model

1. Introduction

Disaster risk is the potential disaster losses in lives, health status, livelihoods, assets and services which could occur to a particular community or society over some

[1] Department of Computer Science, College of Computing and Information Sciences, Makerere University, P.O. Box 7062, Kampala, Uganda.E-mail: gokori@gmail.com
[2] School of Forestry, Geographical and Environmental Sciences Makerere University ,P.O. Box 7062 Kampala, Uganda.E-mail: jobua09@gmail.com.

specified future time period (UNISDR, 2009). A disaster loss results into human activity or condition that may cause loss of life, injury or other health impacts, property damage, social and economic disruption, and environmental damage. International organizations such as the United Nations Office for Disaster Risk Reduction (UNISDR) have facilitated efforts to promote a common understanding of disaster risk reduction concepts and terminology.

This is meant to enhance required knowledge to relate them to socio-economic challenges, sustainable development, environmental management, and humanitarian assistance (UNISDR, 2009). Common terminologies such as hazard, exposure and vulnerability, are frequently used. Hazard exposes people, systems and property present in the hazard zone to potential losses.The losses are due to vulnerability which defines the susceptibility of the community and property to damage effects of a hazard. Hazards may be natural such as earthquake, land slide or man-made such as injury due to a collapsed building, chemical or biological (UNISDR, 2009).

Vulnerability may arise from physical factor which is determined by aspects such as population density levels, settlement, design and materials used in construction; social which refers to the inability of people to withstand adverse impacts to hazards due to characteristics inherent in social interaction and systems of cultural value; economic factors which is dependent on economic status of individuals, communities and nations; and environmental which is due to natural resource depletion and resource degradation (UNISDR, 2009).The probability of harmful consequences, or expected losses such as deaths, injuries and damaged property resulting from interactions between natural or human induced hazards and vulnerable conditions considered as risk has to be assessed continually.

The assessment is to determine the nature and extent of risk by analyzing potential hazards and evaluating existing conditions of vulnerability that together could potentially harm exposed people, property, services, livelihoods and the environment on which they depend. The nature and extent of risk can be modeled with a view of quantifying it to provide an insight into the gravity of a potential risk. Historical data is used as the base input data for the formulation of many risk assessment models therefore, the quality and quantity of such data has a great impact on the accuracy of the risk assessment models. Risk modeling is concerned with an assessment of undesirable events which cause harm and losses for a given time frame and domain of discourse. Risk may be pure risk that features some chance of loss and no chance of gain such as flood, drought or speculative as in the case of investment risk where there may be a chance to either gain or lose (Klosa,2013). This study looks at pure risk with an attempt to design a data model that can be used sufficiently to ascertain levels of risk and minimize losses where it's unavoidable.

Well structured and managed risk database can provide great benefits during preparedness, mitigation, response and recovery when information it holds can be timely retrieved, properly interpreted, and effectively disseminated (Goodchild, 2006). During a disaster strike, data needed by emergency responders may include datasets on roads, utility networks, hospitals, fire stations, damaged areas and damaged facilities since they are vital in rescue operations. Spatial and temporal aspects of disaster events should also be captured in a disaster database for emergency response as it supports assessment of disaster impact (Liu, 2005).

Modeling of risk databases focuses on what data is required and how it is organized. Data modeling is a representation of data structure that includes data objects, association between the objects and the rules that govern operation of the objects (Siricharown, 2009). A data model can be created using an Entity-Relationship or Object-Model approach. This study looked at the Entity-Relationship (ER) approach since it takes less time to design, provides better quality of design and the design is easy to understand (Shoval and Shiran, 1997). A well designed data model is a precursor to a simpler functional model used to design queries to access information from the database and therefore, increases performance of the database.

This study looked at data modeling for a risk database providing the conceptual, logical and physical design models. The conceptual model is used to explore domain concepts and provide initial requirements, logical data model describes the logical entity types, data attributes describing those entities, and the relationships between the entities while the physical data model used to design the internal schema of a database with data tables, the data columns of those tables, and the relationships between the tables.

The data requirements were defined and analyzed using data modeling process to identify needed information which supports risk modeling. The model components identified to be vital in design of risk models are hazard, exposure, vulnerability and loss.

2 Materials and Method

Agile modeling approach coupled with theoretical research based on documents related to risk studies, data storage and information retrieval for disaster scenarios was applied. Continuous improvement was adopted in the identification of all processes, entities, relationships and data flows. The approach was done to the following stages:

(i) Conceptual data modeling
(ii) Logical data modeling
(iii) Physical data modeling

2.1 Model Design

The information which supports risk modeling was derived from the likely questions posed during disaster occurrences and projections. This was used to identify the dimensions used in the modeling and the data modeling process as conceptual, logical and physical model design adopted. The risk database design incorporated the model components which were considered as dimensions in addition to time and location of risks in the model.

Conceptual model design was used to explore domain concepts and how they are interrelated while logical design was used to define the attributes in each of the entities and the relationships between the different entities and attributes. The final step was the physical model design which involved the actual design of the database where objects were defined at the schema level. Objects such as tables and columns were created based on entities and attributes that were defined during the logical modeling.

3 Results

The lead questions posed during an emergency, disaster strike or epidemic outbreaks are as shown in Table 3.1

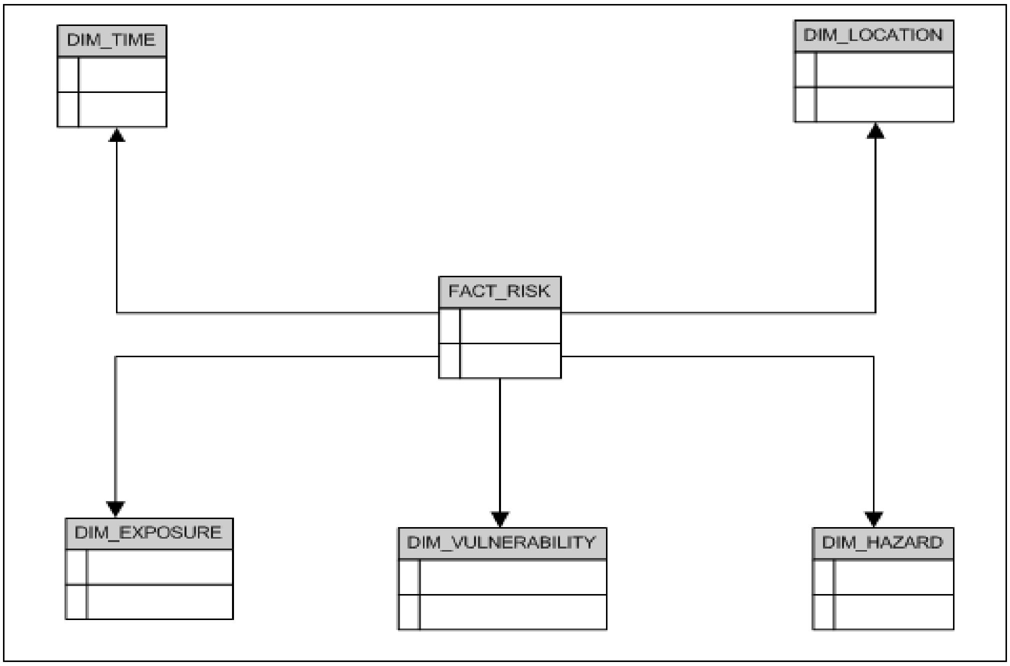
Table 3.1: Lead questions posed during an emergency, disaster strike or an epidemic outbreak

Likely Question	Answer Guide	Measure	Dimension
1. What has happened?	Disaster Type		Hazard
2. Where?	Location		Location
3. When?	1. Date	1. Day 2. Week 3. Month 4. Year	Time
	2. Time	Time stamp	Time
4. What's the magnitude?	1.Death	1. No. of people dead 2. No. of animals dead	Loss
	2. Injury	1. No. of people injured 2. No. of animals injured	Loss
	3.Loss of property	1. No. of buildings destroyed 2. Acreage of crops destroyed 3. Roads blocked 4. Interruption to power supply	Loss
5. What assistance has been provided?	1. Food stuff 2. Medicine 3. Shelter	1. Quantity of the humanitarian assistance	Loss
6. Who provided the assistance?	1. Government 2. NGOs 3. International organizations	1. Cost of basic needs provided	Loss
7. How can the site be reached and how far is it?	1. Road 2. Rail 3. Air	1. How far is the site	Distance
8. What infrastructure is available at the site?	1. Schools 2. Health facility 3. Good roads	1. How many 2. How good is the road, rail and air service	Vulnerability
9. What's the population of the area?		1. No. of people in area 2. No. of Animals	Exposure
		3. No. people aged < 18 years 4. No. of Women 5. No. Disable persons	Vulnerability
10. What's the number of buildings in the area?		No. of buildings	Exposure

The answers to the questions and the likely expectations from the affected community, the respondents and the policy makers guided the derivation of the components synonymous to the dimensions used in modeling. Vulnerability in this study is regarded as a predicative quality used to imagine what could happen to a section of the population in case of certain risks and dangers while exposure is that which is affected by natural disaster such as people and property.

The model components constituted what should be included in a disaster risk model to determine the magnitude of a disaster. In this study, disaster risk is considered to be a function of hazard, exposure, vulnerability and loss. The conceptual model design representing a disaster scenario is represented in Figure 3.1. It visually represents a proposed design of a disaster risk database that associates database objects such as entities with disaster scenarios to ensure completeness and accuracy in requirements gathering.

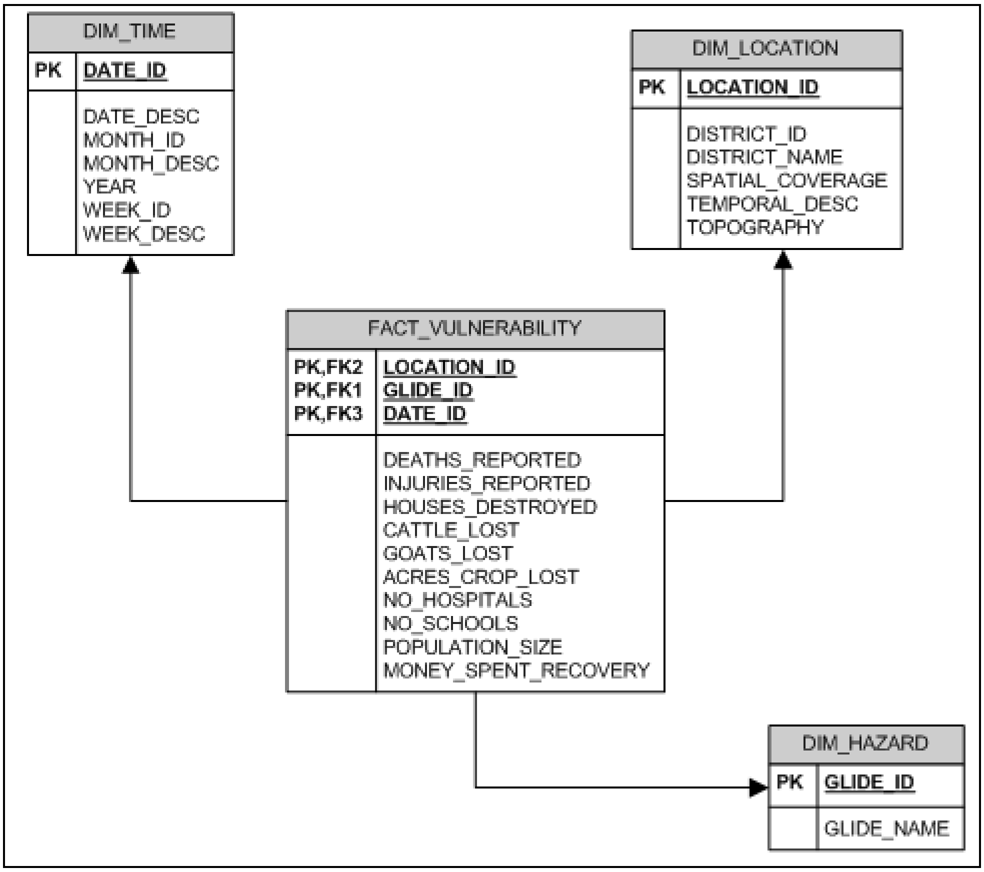
Figure 3.1 Conceptual Model Design



The design scenario indicates that for a hazard such as flood, landslide and earthquake to be disastrous there must be an exposure, vulnerability in a location at a space of time which increases risk expectation that includes value of losses. With information such as time, hazard and vulnerability available we can model the risk to determine the extent of damage of a disaster.

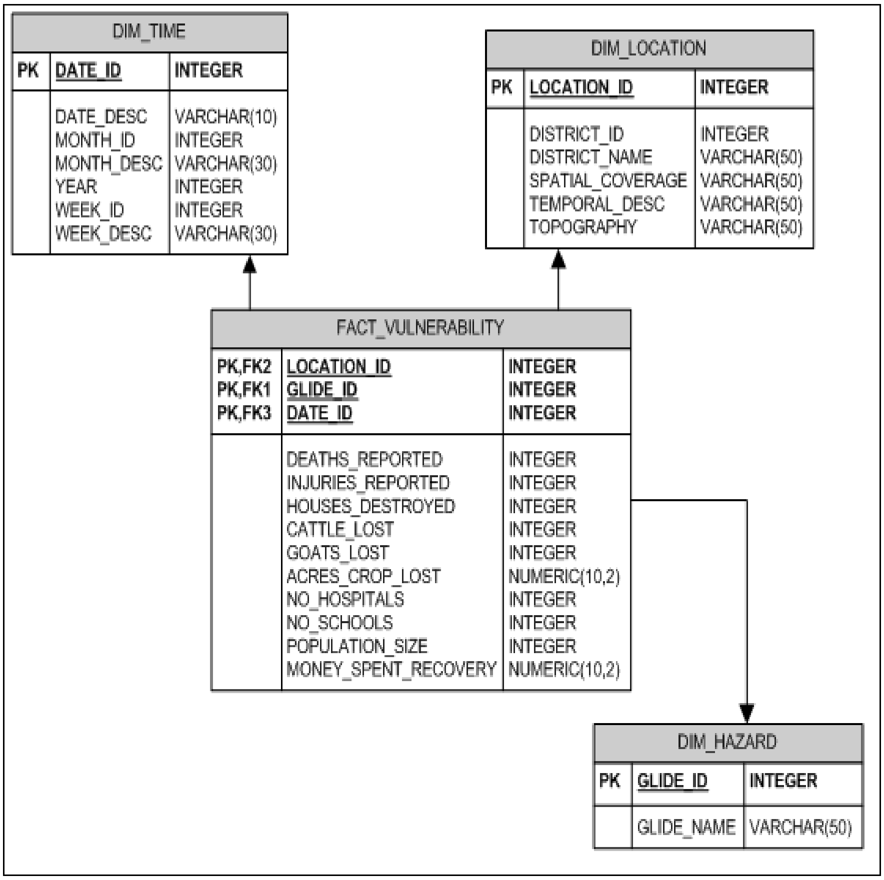
To capture the concept in a database for historical information retrieval and risk modeling purpose the entities in the dimensions such as time, location and hazard are transformed into an entity relationship (ER) model represented as a logical model shown in Figure 3.2. This modeling does not fully support database normalization to the highest normal forms since performance of an implemented database while retrieving information for accurate and quick decision is a major concern. To support this notion, dimensions such as exposure, vulnerability and risk have been collapsed into vulnerability dimension in the logical and physical designs. The fields may be extended to include any relevant data as may be required for any particular disaster.

Figure 3.2 Logical Model Design



;It is evident that a lot more time need to be invested in performing logical modeling so that better options become available for planning the design of the physical database. The physical model design shown in Figure 3.3 represents the actual design of the database and its software specific. The varchar length defined is an example and can be modified accordingly.

Figure 3.3 Physical Model Design



In the physical design (Figure 3.3) objects such as tables and columns are created based on entities and attributes that were defined during logical modeling and constraints defined, including primary keys, foreign keys, and other unique keys. Objects such as indexes and snapshots are defined during this modeling stage. To retrieve information from the designed tables, suitable predicates such as Equal, the joins such as inner join and outer join are used while building structured queries. The nature of a query determines how fast the required information is retrieved from the database.

4. Discussions

Disaster risk management which has stages like risk identification, preparedness, mitigation, response and recovery faces challenges of variability in terms and models applied in each of the stages by various stake holders.

The variability may affect timely intervention to save lives and property during a disaster strike due to ambiguity and misunderstanding in terms used and applied in risk management (IFRC,2013).Harmonization of terms and models used by different stake holders in disaster management improves utilization of information generated, and increases data quality and consistency. It also ensures that the key points are conveyed consistently, even when they are conveyed to different audiences or by different stakeholders (IFRC,2013). As mentioned in FAO (2002), harmonization is a challenge as all needs may not be accommodated in one definition. It is however, more feasible to make the various definitions comparable so that collected data can be easily processed to serve various purposes. This study asserts that harmonization of terms and model components used by stake holders in risk management is vital in enhancing timely intervention and control of disaster risks.

Many areas of the world such as sub-Saharan Africa are struck by disasters resulting from occurrence of eminent natural hazards such as drought, flood, landslide and man-made such as civil war which adversely affect lives and property. Governments too are faced with high cost of redevelopment and hindered sustainable development due to disaster occurrences.

The risk involved depends on the exposure and vulnerability of the communities in the affected areas and therefore should be mitigated to reduce loss of lives and property. As suggested by Mansourian et al.(2006) risk mitigation measures may include financial, planning and legislative measures to avoid possible interactions between hazards and the elements at risk wherever feasible to enhance the resilience of the elements at risk or change the environment when and where the interaction are unavoidable.

Mitigation measures should be supported by readily available, precise and accurate information about a particular risk. Availability of appropriate tools and technologies to facilitate management of disasters is therefore a priority. Performance of the tools and technology is highly depended on completeness and accuracy of collected information and efficiency of information retrieval by users an area this study has elaborated. As argued by Mansourian et al. (2006), Liu (2005) and Chen (2004) efficient and effective sharing of information on disasters by different users depends a lot on representation, organization and accesses to relevant information on a particular disaster. This can be realized through appropriate data modeling which serves as a link between disaster information needs and system requirements. In this study, data modeling concept which emphasizes representation and documentation of requirements needed to support risk modeling processes during application software design and development has been applied with success. The study has demonstrated that a properly design risk database is an impetus for enormous progress in information sharing in risk related scenarios.

Data modeling of risk databases follows the modeling process which begins with requirements capture to actual design of the database (Letkowski,2014; Meredith, 2007) . The process which produces conceptual, logical and physical data model respectively is explored in this study with the ultimate goal to explore risk domain concepts up-to design of internal schema of a risk database.

How the data is stored determines the performance of the database and it manifests itself into the speed at which data is retrieved from a particular database. This study suggested a logical model design which can be used to provide an appropriate relationship between different tables in a risk database with consolidated, transactional and organized data. The model facilitates querying of data in in a dimensional structure with a staging space for analysis of information

In-order that a compromise between performance of the database and completeness of information needed to support decision in a disaster scenario is attained; some entities like exposure and vulnerability are collapsed in a risk dimension. It is envisaged that such operation of combining dimensions when carefully done without losing details of information can improve performance of a database making information retrieval faster to support timely intervention in a disaster scenario or mitigate a risk on time.

5 Conclusions and Recommendations

The components regarded as dimensions required in the design of a risk model were identified as hazard, vulnerability, exposure and loss. The dimensions were used in the conceptual, logical and physical model design of the risk database and this supports timely retrieval of information to mitigate disaster impact. Cross-sectoral integration of information was applied to each of the stages of modeling to enhance completeness in the design.

It is recommended that there should be integration within each of the elements such as terms, models and databases to ensure that counties are well equipped with high quality data and information to anticipate and avert disasters such as flood, drought and landslides. The data model should be well structured with the aim of improving performance of the system to enhance timely information retrieval.

References

Chen, Y., 2004. GIS and remote sensing in hydrology, water resources and environment, Wallingford, International Association of Hydrological Science.

Food and Agriculture Organization, 2002. Harmonizing forest-related definitions for use by various stakeholders. FAO Rome.

Goodchild, M. F., 2006. GIS and disasters: Planning for catastrophe (Editorial), Computers, Environmental and Urban Systems, 30, 227-229.

International Federation of Red Cross, 2013. Public awareness and public education for disaster risk reduction, Geneva, Switzerland.

Klosa E., 2013. A concept of models for supply chain speculative risk analysis and management, Journal of Economics and Management, 12, 46-59.

Letkowski J., 2014. Doing database design with MySQL, Journal of Technology Research

Liu G., J., 2005. A report on EMA disasters database enhancement. RMIT University (2005).

Mansourian, A., Rajabifard, A., Zoej, M. J. V. &Williamson, 1., 2006. Using SDI and web-based system to facilitate disaster management, Computers & Geosciences, 32, 303-315.

Meredith M,. 2007. Data modeling a process for pattern induction, Journal of Experimental & Theoretical Artificial Intelligence, 3 (1) 43-68.

Shoval, P. and Shiran, S., 1997. Entity-relationship and Object-oriented data modeling – an experimental comparison of design quality, Data and Knowledge Engineering, 21, 297-315.

Siricharoen, W. V., 2009. Ontology Modeling and Object Modeling in Software Engineering, International Journal of Software Engineering and Its Applications, 3(1), 43-60.

Umanath N. S. and Scamell R. W., 2007. Data Modeling and Database Design, Thomsom Learning Inc, Boston,MA, USA.

UNISDR, 2009. Terminology on Disaster Risk Reduction, Geneva, Switzerland.

West,M. andFowler,J., 1996. Developing High Quality Data Models. The European Process Industries STEP Technical Liaison Executive (EPISTLE).