Use of Incident Data Collection from Various Sources For Industrial Safety Performance Assessments

Nir Keren, T. Michael O'Connor, and M. Sam Mannan[1]
Mary Kay O'Connor Process Safety Center, Chemical Engineering Department, Texas A&M University System, College Station, Texas 77843-3122, USA
Phone: (979) 862-3985, Fax: (979) 845-6446, mannan@tamu.edu

Keywords: incident database, safety performance assessment, data collection, incident trend analysis

Abstract

Many organizations collect data on industrial incidents. These organizations differ from each other in their interests, data collection procedures, definitions, and scope, and each of them is analyzing its data to achieve its goal and to accomplish its mission. However, there were no attempts to explore the potential hidden in integrating data sources. Extensive efforts are required in order to integrate information from different data sources as well as to identify the effects of the individual aspects of data collection procedures on the quality and completeness of the data. This paper describes a methodology for incident data collection from various sources, and the opportunity that exists in a combined data mart for industrial safety performance assessment and identification of trends. Additionally, such analysis can be used to determine the areas for major reduction of losses and reduction in the number of incidents.

Introduction

There is an increased interest in using data on accidents to improve safety in the last 20 years. At the late 80s, V. C. Marshal consolidated incident data from sixty or so years and harnessed it toward loss reduction, and loss prevention in his book "Major Chemical Hazards" [1]. Today the interest is bigger than ever, because of the development of information technologies that looks promising in their abilities to see what "unarmed human eye" cannot see. Major efforts are being invested toward collection of incident related data. The US Department of Health and Human Services, The Agency for Toxic Substances and Disease Registry (ATSDR) maintain hazardous substances emergency events surveillance (HSEES) and publishes annual and cumulative reports [2], and is only one among many other type of data collection projects that is maintained by the Center of disease Control (CDC). The Department of transportation repository consist of large number of transportation safety Related databases, and many reports are available on their website [3]. The last are only two from at least 15 sources of information of incident related data that have been analyzed and incorporated in assessments of industrial safety performance by the Mary Kay O'Connor Process Safety Center, at the Texas A&M University, College Station Texas (MKOPSC). However, the main challenge in using incident related data only begins when the data is available.

Marono et al. suggests to use the European Commission accident-reporting database MARS as a support for the definition of a safety performance indicator system [4]. McCray and Mannan are the first to look in several databases for opportunities for risk reduction and loss prevention [5]. Mannan with O'Connor and West established the basis for a continual effort to exhaust the potential that is hidden in accident databases in their paper in reference 6. Mannan et al. looked again into EPA RMP Info database in order to determine the most significant chemical releases

[1] Corresponding author

[7] as part of the efforts described above. Early at 2002 the MKOPSC established a report on the feasibility of using federal incident databases to measure and improve chemical safety [8], and another report on assessment of the chemical safety in the United state for 2001 [9].

This paper presents the methodology that is being used by MKOPSC, the challenges, difficulties, measure, and shortly discusses future research and development to improve this methodology and increase its quality and capacity.

Assessment that is based on a methodology of incident data collection from various sources is a thorough process that has to be done carefully and in several stages. The flow chart in figure 1 is a simplified description of the process. The primary focus of industrial safety performance assessment, which uses the methodology described herein, is to establish a baseline metrics for the universe under investigation with regard to safety. This requires identification of incident trends, distribution of number of incidents, number of injuries, property damage costs, releases of materials, hospitalizations, and evacuations. These should be analyzed and correlated across the causes of incidents, equipment involved, initiation events, location, and other indicators. Several of the sources that are available collect only part or a sample of the information. However, it is possible to estimate the total number of chemical/product related incidents by applying statistical tools on the data. Implementation of indicator-based industrial performance measurement systems helps to determine whether the efforts invested toward safety improvement lead to the desired results. Other benefits are the ability to determine the areas that will lead to major reduction of losses and reduction in the number of incidents.

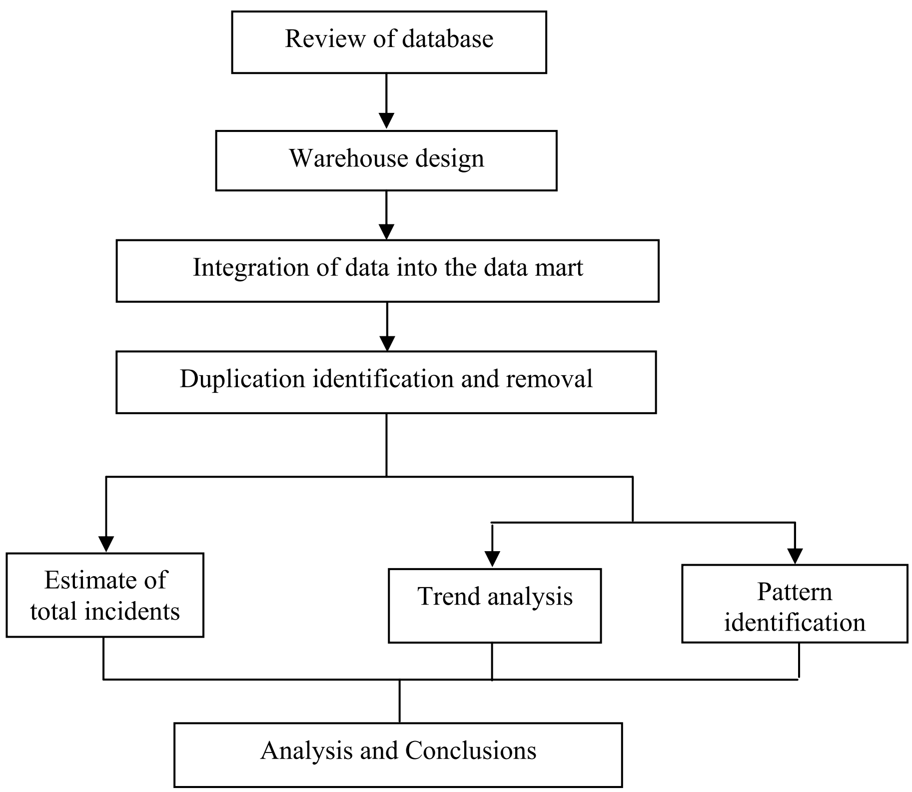

Figure 1- Methodology Flow Chart

Among the major conclusions from studies that have been conducted using this methodology is to not be "misled" by the amount of data that a certain source may consist of. In one study, a source of information provided about two-third of the data; however, it failed to collect significant data (e.g., failed to collect data with severe consequences). This conclusion justified the efforts that were required to broaden the search and combination of sources of information. A novel data collection methodology, based on News Clippings, has been established by the Mary Kay O'Connor Process Safety Center. This method uses search engines to query newspapers according to a predetermined set of keywords. The information is collected and submitted to the datamart. This method has several advantages including the ability to further investigate the incident or to verify the information if required.

Sources of Information

The process of integration of data from several sources requires a thorough analysis of the databases that collect information on industrial incidents. Table 1 consists of a list of more than a dozen databases from ten sources that were integrated for an assessment project for a certain industry sector. These databases were selected because they contain information that could be used to establish safety performance metrics for the industry sector.

The form of the data reflects the interest, purpose, and scope of the organization collecting the data. The lack of national and international standard of reporting incidents as Johnson mention in reference [11] has led to a lack of consistency among the sources with regard to coding used in the variety of fields. As a result, major efforts are needed to create an infrastructure that will allow data from variety of sources to "sit" together in a datamart. Figure 3 demonstrates the information flow until it reaches its final destination. At almost every node the data is being converted, and the process must be done diligently in order to avoid misinterpretation of the data.

It is important to emphasize that the sources do not release the information as it becomes available. A real-time data collection from various sources is a long process that takes at least three years, as can be seen in Figure 2. Because of its real-time nature, the news clipping data collection system creates several opportunities:

Table 1 - Sources of Information and Databases

Source	**Database**
Federal Emergency Management Agency (FEMA)	National Fire Information Reporting System (NFIRS)
U.S. Consumer Product Safety Commission (CPSC)	* National Electronic Injury Surveillance System (NEISS) * Death Certificates * Investigation Summary * Incident Summary
Mary Kay O'Connor Process Safety Center (MKOPSC)	News Clipping Database
States Associations	State of Iowa State of Florida
State Agencies	State of Texas
National Response Center (NRC)	Incident Reporting Information System (IRIS)
US Department of Health and Human Services, Agency for Toxic Substances	Hazardous Substances Emergency Events Surveillance (HSEES)

and Disease Registry	
U.S. Department of Transportation (DOT)	* Hazardous Material Incident Reporting System (HMIRS) * Integrated Pipeline Information System (IPIS) also known as Hazardous Liquid Accident Data (HLAD).
U.S. Environmental Protection Agency (EPA)	*Risk Management Program (RMP) 5-year Accident History *Accidental Release Information Program (ARIP)
U.S. Department of Labor, Occupational Safety and Health Administration (OSHA)	Accident Investigation System and several other databases.

- Development of procedures for incident investigation for the real-time data collection
- Identification of need for incident investigation and performing investigation
- Follow-up on information to validate causes of incidents and long-term consequences

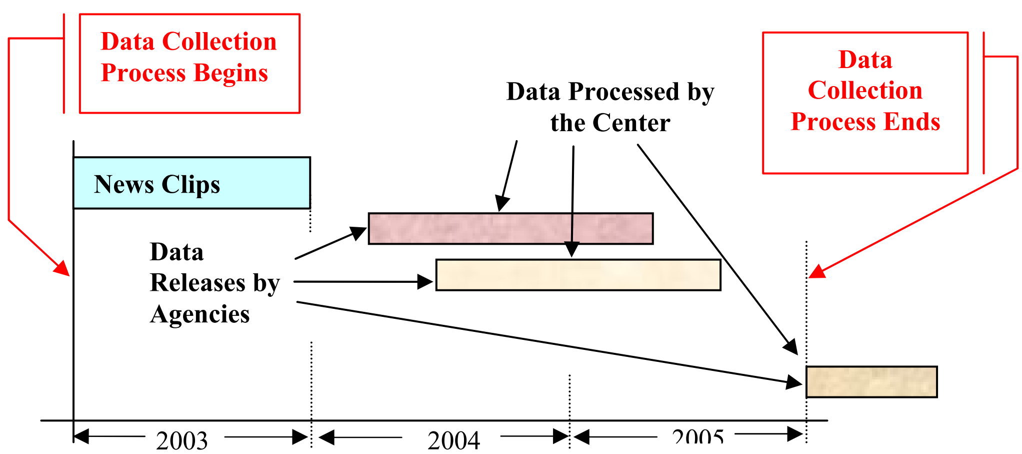

Figure 2 - Timetable of Real-Time Data Collection and Analysis

In many of the records that were examined, question arose that could have been answered quite easily if a real-time data collection process had been in place. In several of the records, it was hard to determine what is the cause, or what was the initiating event. As an example, one of the records contained data for an incident in Alaska.

The record indicated 99 fatalities for the incident. Since it is reasonable to assume that an incident with such large number of fatalities would be covered by the media as well as by incident investigation reports, a thorough research was conducted, which revealed that the incident actually resulted in a single fatality and 99 injuries.

<u>Method of Duplication Identification and Removal</u>

At the end of the data submission stage, it is required to identify duplications and to remove them. Johnson discusses many of the problems involved in automation of the process of duplication identification in reference [11].

There are two categories of duplications that are encountered during the consolidation of incident

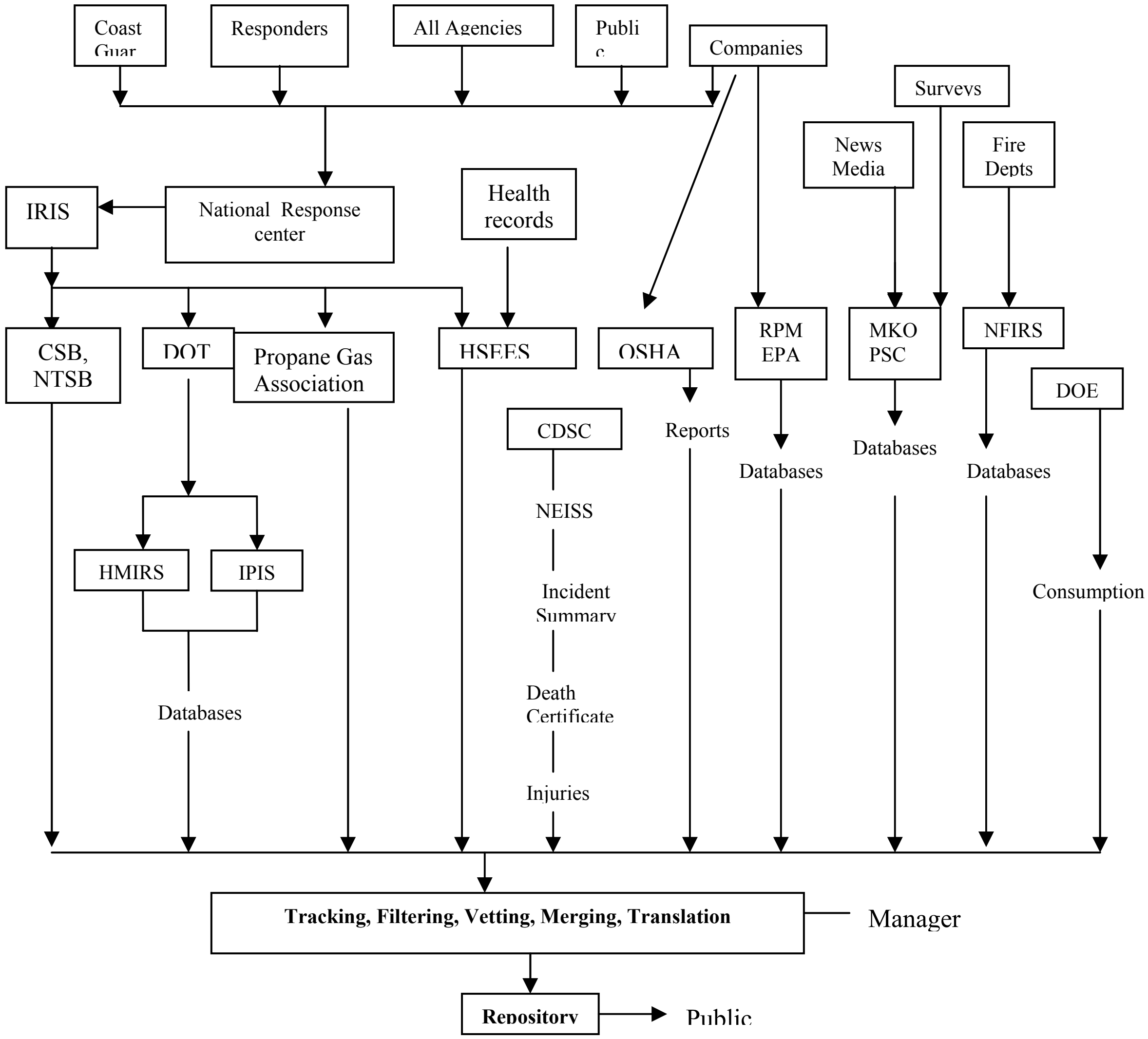

Figure 3 - Information Flow Chart

of incident information from a variety of sources:

- Duplications within the sources themselves, and
- Duplications among different sources

In general, it is much easier to identify duplications within the sources as compared to identifying duplications amongst different sources. However, the process of identification of duplications is similar in both cases. Duplication within the same source has the same type of information and is much easier to identify. The duplication identification process is illustrated in figure 4. The number of records in the list of 'Suspected as Duplications' is sensitive to the time frame that is employed. However, in order to verify that the time frame used is not arbitrary, the Center studied the sensitivity of the number of suspected as Duplications to the time frame.

As Figure 5 reveals, the number of incidents that are suspected as duplications is highly correlated with the width of the time frame (root mean square value of more then 0.98).

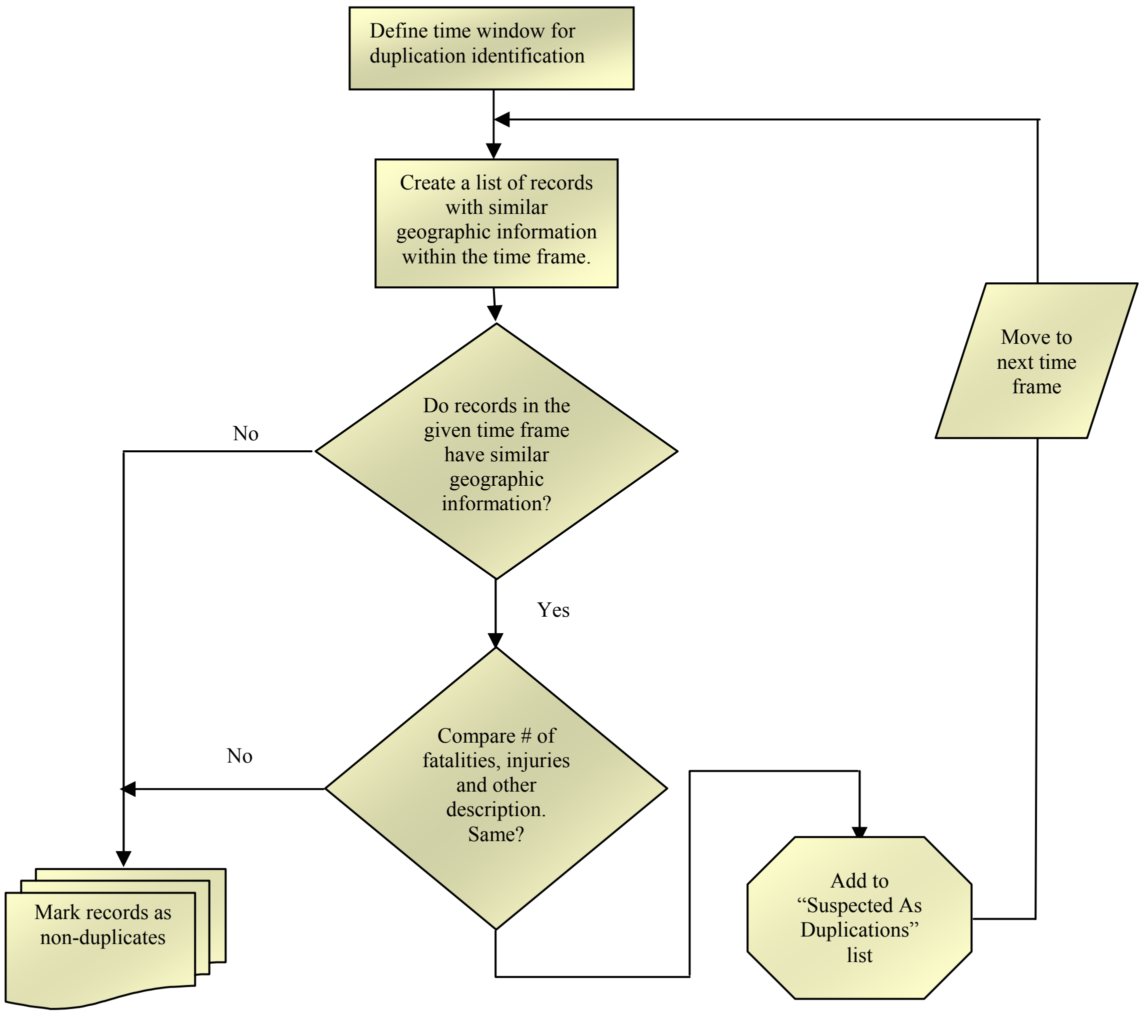

Figure 4 - Procedure for Identification of Duplications

The slope of the correlated line can serve as a qualitative relative indicator for the comprehensiveness of the database. Under the estimation that the probability of an incident to occur is not time dependent, the number of suspected duplication in a given time frame would increase as the portion of the universe of incidents increases. The slope of the curve becomes steeper as the comprehensiveness of the database increases.

Once the system creates a list of records that are suspected as duplications, the records are reviewed manually, and a decision with regard to these records are made. Records that are identified as duplications are marked, so queries will reveal only one of them. Identification of duplicates becomes quite difficult in cases where time of incident is not given.

As for duplicate identification within the databases, the process of verification of whether incidents are duplications varies according to characteristics of the incidents. NFIRS for example contains two types of duplications:

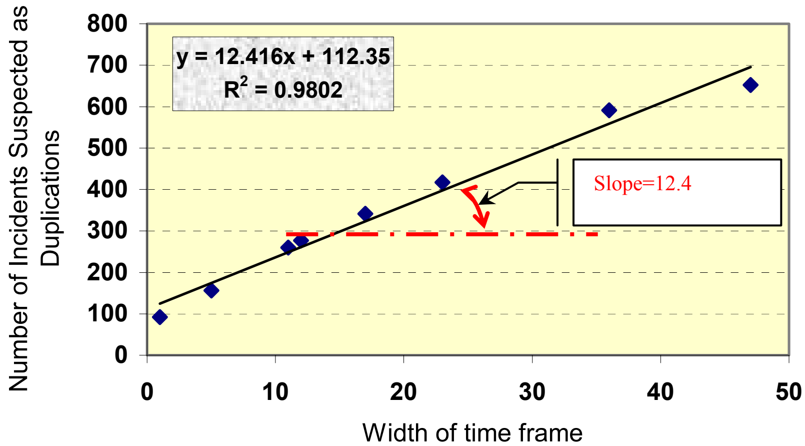

Figure 5 - Sensitivity to Time Frame Study

1) Fire department that reported the same incident more than once.
2) Incidents that were suspected as duplicates, because more than a single fire department reported the incident.

In the first case, the verification process is not complicated. In the second case, however, it was necessary to conduct an Internet search for county maps in order to determine if it is reasonable that a fire department from an adjacent county would assist another fire department and also report to NFIRS. In the majority of the cases the distance between the counties was too far to assume that the reports are duplicates.

An important criterion for identifying duplications is the number of injuries and fatalities. If two incidents that have other similar characteristics also show exactly the same number of fatalities and injuries, there is a high likelihood that one of these incidents is a duplicate. The system ignored incidents that have different number of injuries or fatalities. The Center applied manual checks and quality control procedures to ensure that duplicates were identified accurately and that non-duplicates were not eliminated inadvertently.

As for duplications amongst different databases, the process required relatively more extensive efforts, and each of the cases needed to be treated separately.

<u>Methodology for Estimation of Total Number of Incidents</u>

The process for estimating the total number of Industrial incidents in the Unites States can be explained by the theory of sets. Figure 6 illustrates the current situation. The gray area represents the total number of Industrial related incidents in the US. The white areas represent the actual number of incidents in each of the respective databases.

The number of incidents from each of the databases is a subset of the total number of incidents that this database would consist of if all incidents were reported to the source (the set). For example, NFIRS, which is a database that consists of reports from emergency departments, contains records from about 14,000 fire departments from 42 states. The records in NFIRS are a subset of a set, which is the number of records that NFIRS would consist of if all 29,000 fire departments as well 6,900 emergency departments from the 50 states reported every Industrial incident to NFIRS. Figure 7 is an illustration of the relation between a set and a subset.

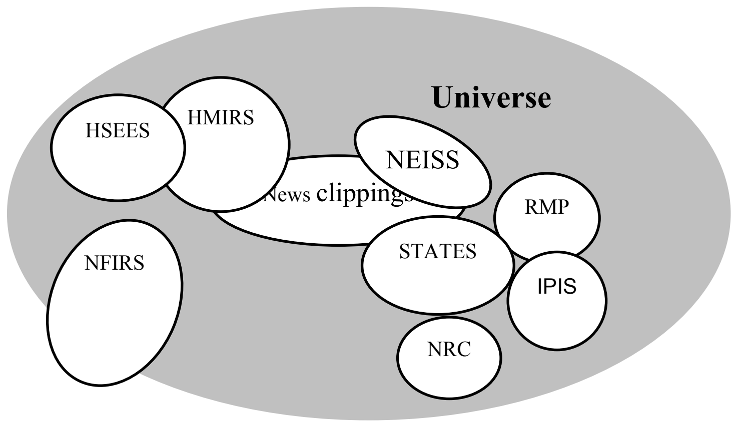

Figure 6 - Illustration of relation between total number of Incidents and Number of Incidents in the Sources

The Universe is a collection of all incidents that have the potential to be reported. Therefore, Universe is a composition of sets. The translation of the above to the theory of set language is as follows:

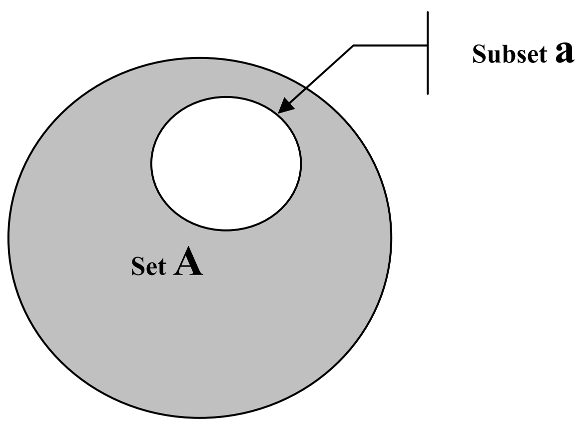

Figure 7 - Relation Between a Set and a Subset

$\mathbf{a_1}$ **-** is current records in database **DB1**
$\mathbf{A_1}$ **-** is the potential number of record in the database DB1, if all incidents targeted by this database were reported.

$\mathbf{a_1}$ is a subset of $\mathbf{A_1}$ ➔ $\mathbf{a_1} \subset \mathbf{A_1}$

$\mathbf{a_2}$ **-** is current records in database **DB2**
$\mathbf{A_2}$ **-** is the potential number of record in the database **DB2**, if all incidents targeted by this database were reported.

a_2 is a subset of **A_2** ➔ **$a_2 \subset A_2$**

The same principles applies to $a_3, a_4, \ldots\ldots, a_n$ or all the databases.

The Universe **S** is a composition of all the sets. However, there are overlaps among the sets, and therefore U is a union of the sets, as equation 1 shows:

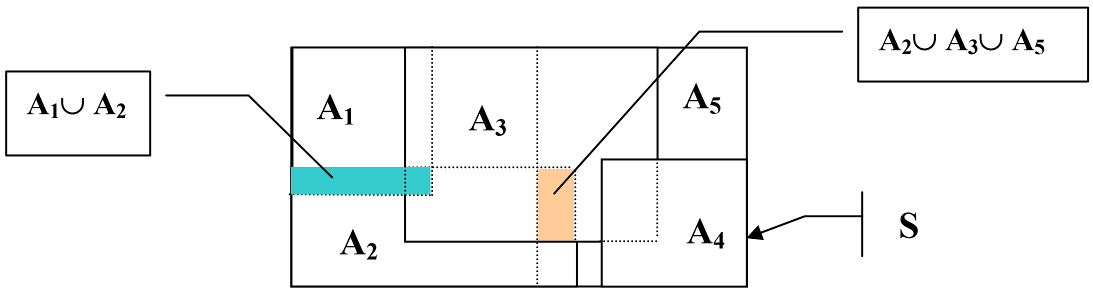

Figure 8 - The Universe is a Union of the Sets (Venn Diagram)

$$S = \left(\bigcup_{i=1}^{n} A_i\right) = A_1 \bigcup A_2 \bigcup A_3 \cdots \bigcup A_n = \overbrace{\sum_{i=1}^{n} A_i}^{\text{The sum of incidents from all databases prior to applying duplication identification procedure}} - \overbrace{\sum_{i=1}^{n} \sum_{j>i}^{n-1} \left(A_i \bigcap A_j\right)}^{\text{The sum of the number of duplicates between every combination of pairs of databases}} - \tag{1}$$

$$- \underbrace{\sum_{i=1}^{n-1} \sum_{j>i}^{n-2} \sum_{k>j}^{n-3} \left(A_i \bigcap A_j \bigcap A_k\right)}_{\text{The sum of the number of multiplications among every combination of three databases}} - \ldots - \underbrace{\left(A_1 \bigcap A_2 \bigcap \cdots \bigcap A_n\right)}_{\text{The number of multiplication that appeared in all of the databases}}$$

No multiplications found between more than two sources. Therefore, only the first two parts of equation 1 will be employed for the estimation purposes. These two parts are extended and are presented in equation 2:

$$\mathbf{S = A_1 + A_2 + A_3 + \ldots . + A_n - [(A_1 \cap A_2 + A_1 \cap A_3 + \ldots + A_1 \cap A_n) + + (A_2 \cap A_3 + A_2 \cap A_4 + \ldots + A_2 \cap A_n)] + \ldots + (A_{(n-1)} \cap A_n)]} \tag{2}$$

The sequence of estimating the universe S is now simplified. The information that is available currently is the subsets a_i and the intersection between these subsets. Figure 9 presents the sequence of obtaining the information required to solve equation 2.

Following figure 9 is a description of the process for extrapolating the sets $\mathbf{A_i}$ according to the characteristics of each of the sources. The assumptions that were required in order to extrapolate the intersections between the sets are presented later.

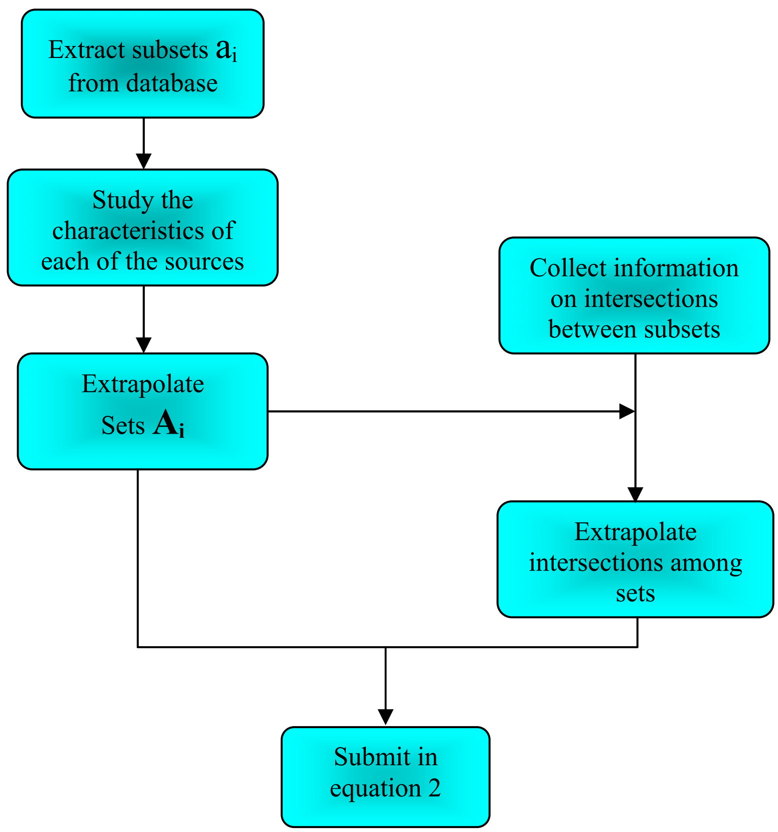

Figure 9 - Sequence of Estimation of Universe S

Extraplation of Sets A_i: The purpose of collection of information is not the same for all the sources, and therefore the characteristics of each of these sources should be incorporated in order to calculate the number of incidents that the source database would consist of if it were to capture all the incidents that belong in its category. A set of considerations, as well as the methods for extrapolating of information from the sets, A_i, is developed individually.

Extrapolation of Duplicates: The ideal way to extrapolate the number of duplications is to sample several sample sizes of sub-sets and to identify number of duplicates for combination of sizes. By using this methodology it is possible to study how the number of duplications increases with increase of the size of subsets. However, in cases where the databases consist of relatively low number of duplicates, an approximation can be done by multiplication of the number of duplicates between sources by the ratio of the sum of the extrapolated number of the incidents and the sum of the actual number of incidents in the database.

Calibration of Information from Sources

When reviewing sources it is important to verify that the data represents the relevant population uniformly. The Consumer Product Safety Commission (CPSC) - National Electronic Incidents Surveillance System (NEISS) is a collection of injury data that are gathered from the emergency departments of 100 hospitals selected as a statistical sample of all 5,300 U.S. hospitals with emergency departments. NEISS surveys sample of hospitals that represent all ethnic groups and concentrations of population, and it is statistically valid to extrapolate this data. However, the Federal Emergency Management Agency, National Fire Information Reporting System (NFIRS) database, is a collection of incident reports from fire departments. About 30% of the fire departments in the US, from 42 states are report incidents to NFIRS. These fire departments vary in size range from departments that protect several dozen individuals (rural areas) up to departments that protect millions. In large fire departments all employees are paid, and in small fire departments employees are volunteers. The probability that large fire departments will report to NFIRS is much higher than small fire departments. Therefore, the analyst using NFIRS as a source of data should be aware of the distribution of consumption of the product that is under investigation. If fertilizers are the product in study, then the distribution of consumption in urban areas is expected to be much lower than the consumption in rural areas. However, rural fire departments in rural areas mainly employ volunteers and therefore the probability that these fire departments will report to NFIRS is low. In that case, information from NFIRS may be biased, and a calibration should be conducted. The Center used a survey of fire departments for calibration purposes.

Use of Indicators Toward Industrial Safety Performance Assessment

What are Indicators and What Do They Mean: As noted previously, a large amount of information exists about industrial incidents, including a large amount of information gathered by federal, state, and local agencies. The information gathered includes data on the specifics and numbers of releases of chemicals, on injuries, illnesses, deaths caused by chemicals and other products. Are any of these though accurate indicators of the state of effectiveness of chemical safety efforts? Do they tell us whether we are making progress in chemical safety?

An indicator is generally defined as an observed variable. Essentially, an indicator is presumed to reflect through a positive correlation a single underlying variable. The underlying variable being considered here is the safety of chemical processes. It is impossible to observe or measure industrial safety as a positive measure. It can only be measured as a negative measure, or an observable variable which is defined as when safety processes fail. The number of failures is an indicator, when taken in the context of the universe of potential failures, of industrial safety.

The indicator becomes more valuable in understanding the underlying variable when looked at over a period of time or as a trend. Trend analysis looks at an indicator or series of indicators over a period of time to observe if there is a general sustained movement of the time series upward, downward, or if there is no discernible pattern. Trend lines are used to graphically display trends in data and to analyze problems of prediction. Such analysis is also called regression analysis. By using regression analysis, it is possible to extend a trend line in a chart beyond the actual data to predict future values. The specific techniques that are most commonly applied include linear model, an exponential model, or a moving-averages model.

Trend analysis is commonly misapplied. For example, two or three data points cannot be used to develop a trend, though under a simple “eyeball” analysis it might seem so. In any trend and regression analysis, there always exists the assumption that a component of the underlying variable is generated through a random or stochastic process interacting with the concrete set of data. Over a short period of time, the potential impact of this random process can be much larger than over a longer time period, where it becomes the "noise" or part of the error term in a regression analysis.

It is often better to use a number of different time periods in completing a trend analysis. For example, weekly measures viewed over a period of a year may indicate an upward movement of the number of injuries related to chemical releases. When viewed over a five-year period, the trend may be generally down, except for the current period, which could have been caused by an external variable such as a change in the definition of an injury, or a change in measuring techniques or methodologies.

On a larger perspective to be able to compare one set of indicators, for example for chlorine, to a set of indicators for petroleum products, the indicators must be normalized so that a comparison is made of essentially equal sets. Normalization is a general process by which two or more indicators are divided by an equivalent denominator. For the above example, an equivalent denominator might be the amount of chemicals produced. It is unadvisable to attempt to make a comparison across indicators that have not been normalized, as there is no actual basis for comparison and the resulting analysis is methodologically indefensible.

<u>The Effect of Policies on Safety in the Chemical Industry</u>: It is as importance to properly select the indicators, as it is to have an idea of what type of information you hope to see. The effects of changes in government regulations covering the chemical industry should be identifiable in the data.

If a specific policy change or new regulation has an effect on industrial safety, then graphic representations of the data recorded in databases would be reflected in the metric of interest. For example, the following graphic might illustrate the results of a governmental policy change. The performance in years one through five is relatively constant. During the fifth year (point A on the chart), a policy change is made and the resulting performance is shown by the value in year six (point B on the chart). It could be inferred the change resulted in about a 40 percent decrease in the number of incidents.

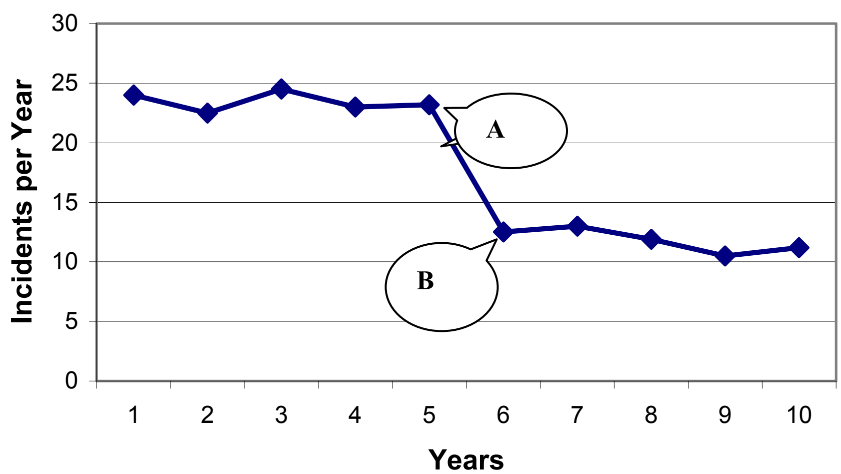

Figure 10 - Measuring the Impact of Policy Changes

Using indicators to investigate databases creates the opportunity to compare performance qualitatively among industry sectors. Figure 11 shows a plot of fatalities recorded by OSHA where a chemical is the primary or secondary cause of the fatality. Figure 12 presents a plot of fatalities from transportation related chemical incident. While figure 11 demonstrates a downward trend in fatalities resulting from chemicals, fatalities from transportation-related chemical incidents shows a slight upward trend over the 10-year period.

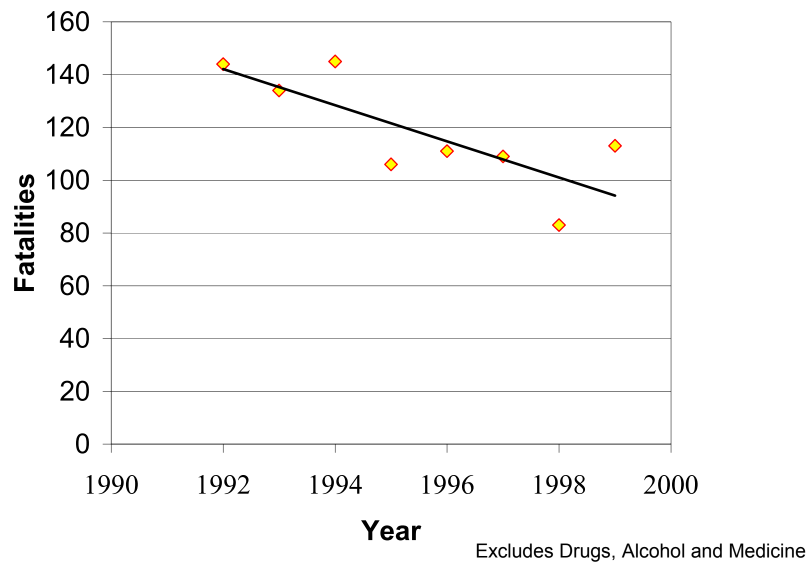

Figure 11 - OSHA Fatalities with Chemicals as Primary or Secondary Source

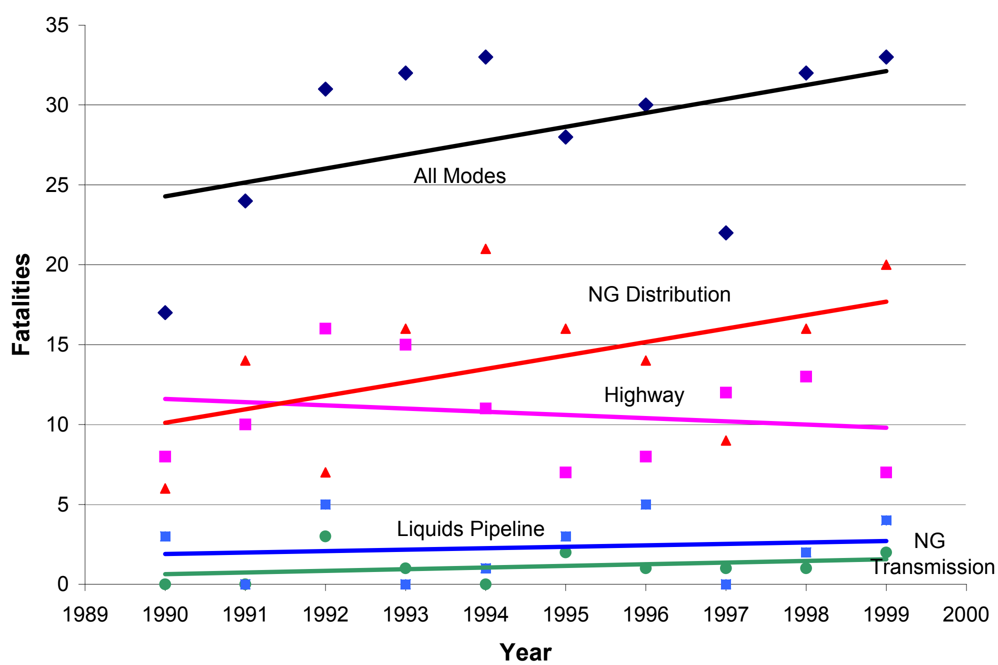

Figure 12 - Fatalities from Transportation-Related Incidents, Due to Chemicals

Figure 13 shows the result of implementation of data collection from various sources methodology on a petrochemical product for a certain year, and figure 14 demonstrates patterns of causes of incidents for the same product.

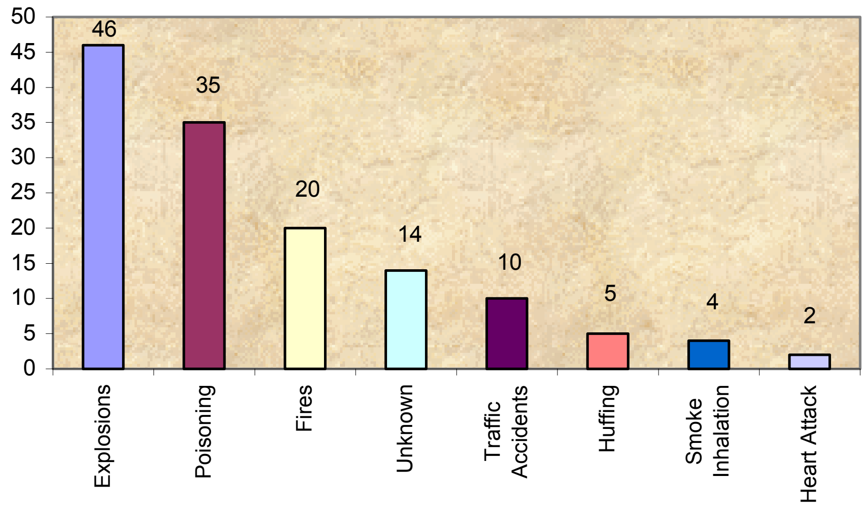

Figure 13 - Distribution of Fatalities by Cause of Death

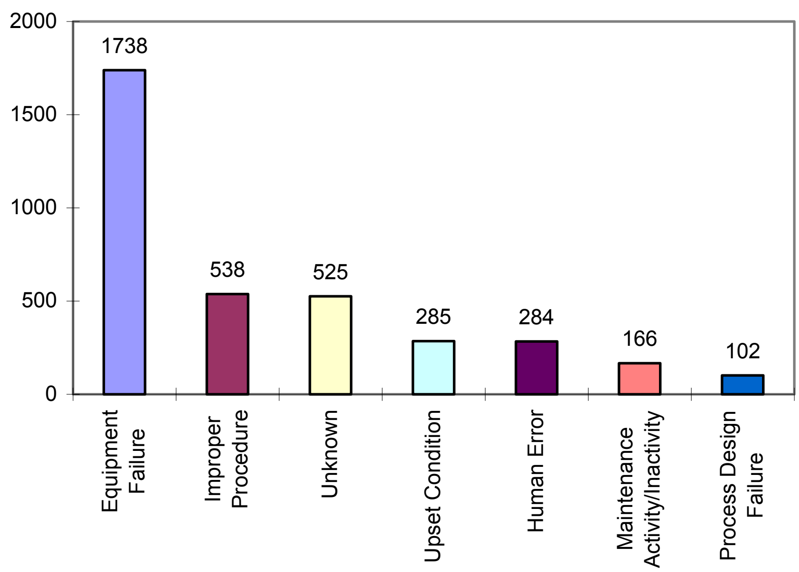

Figure 14 - Distribution of Number of Incidents by Causes

In general, collection of information of product is done in phases. First phase is implementation of the methodology on a single year in order to be able to learn difficulties that are related to collection of data on the product that is under study. After learning and implementing improvements that are required, a second phase is launched where data is collected for several years.

Further R&D

Extensive efforts are required in order to integrate information from the data sources as well as to identify the effects of the individual aspects of data collection procedures on the quality and completeness of the data. A holistic approach that suggests an innovation of tools, methods, techniques, and procedures in order to extract information from the large variety of sources of information is required. The principles as well as the methods of such a holistic approach will be applicable to many other disciplines such as civil engineering and insurance entities. The following is a short summary of the activities that will be conducted in order to automate the methodology that is described in this document: (1) development of methods for the detection and repair of common problems based on the individual storage formats; (2) development of techniques to allow identification and intervention of non-standard problems of these storages; (3) development of methods to automatically identify relational steps based on experts' seed knowledge; (4) merging the techniques above with current cleaning techniques to produce files that could be further processed without concern for storage of format irregularities; (5) development of methods for automatic integration of fields and relation building; (6) conversion of textual information into a schema of warehouse; (7) because of the multi-source legacy data, the development of duplication detection techniques as well as duplication handling techniques will be required as well; (8) allowing generation of flexible user application that prevents the need for involving content expert again, and yet ensure that subject relevance was maintained in the application output.

Conclusions

Several industrial safety performance assessments studies had been done by incident data collection and consolidation from a variety of sources. These studies demonstrated that it is worthwhile to collect data from variety of sources, and that much can be learned from the consolidated database. However, in order to accomplish the ultimate goals of safety performance assessment the consolidated database must include root cause information. In order to do that, a real-time incident data collection procedure must be established. Two major reasons for a real-time data collection process are: (1) news archives make the data available for a short period of time only; (2) the ability to further investigate an incident and get reliable results, decreases significantly with time.

There is enormous potential in employing data collection from a variety of information sources. This technique not only increases the amount of data captured by individual sources but also the ability to capture more diverse and significant incidents. In one of the studies, the methodology used by the Center resulted in the identification of 35% more incidents than the number of the incidents in the single largest source. In the same study, the methodology captured about ten times more fatalities than the single largest source.

The Center applied news clipping data collection procedure as a data collection methodology. However, this methodology is maximized only when applied in real-time because the data sources are available for limited periods of time, and because the Center can solicit additional useful information only during the period shortly after the incidents occur.

Analysis of the data identified a relatively low number of duplications. The majority of the duplications were found within the sources and not between them, i.e., the duplications are mainly because operators reported some incidents twice (or more). We believe that significant improvements can be made by real-time data collection.

References

1. V. C. Marshall, *Major Chemical Hazard,* Halsted Press: a division of John Wiley & Sons, NY, USA, 1987.
2. Division of Health Studies, Epidemiology and Surveillance Branch, Hazardous Substances Emergency Events Surveillance (HSEES): annual reports, http://www.atsdr.cdc.gov/HS/HSEES/
3. Department of Transportation – The Bureau of Transportation Statistics website: http://www.transtats.bts.gov/
4. M. Marono, R. Sola, J. A., J. Santamaria, *Using MARS database as Support for the Definition of a Safety Performance Indicator System,* Proceeding of the 10th International Symposium of Loss Prevention and Safety Promotion in the Process Industries, Stockholm, Sweden, June 2000, pp. 265-272.
5. Eboni McCray, M. Sam Mannan, *Use of Accident Database for Systematic Evaluation of Chemical Accident*, Proceeding of the Mary Kay O'Connor Process Safety Center annual Symposium, College Station, Texas, October 1999, pp. 47-161.
6. M. Sam Mannan, T. Michael O'Connor, Harry H. West, *Accident History Database: An Opportunity"* Environmental Progress (Vol. 18, No. 1), Spring 1999.
7. M. Sam Mannan, William J. Rogers, Fahad Al-Quarashi, *Analysis of the EPA RMP Info Database*, Proceeding of the Mary Kay O'Connor Process Safety Center annual Symposium, College Station, Texas, October 2000, pp. 39-45.
8. The Mary Kay O'Connor Process Safety Center, Texas A&M University, *Feasibility of Using Federal Incident Databases to Measure and Improve Chemical Safety, April 2002.*
9. The Mary Kay O'Connor Process Safety Center, Texas A&M University, *2001 Assessment of Chemical Safety in the United State, April 2002.*
10. Johnson, C.W., ***Using Case-Based Reasoning to Support the Indexing and Retrieval of Incident Reports***, Proceeding of European Safety and Reliability Conference (ESREL 2000): Foresight and Precaution, Balkema, Rotterdam, the Netherlands, 1387-1394, 2000.
11. Johnson, C.W., ***Software Tools to Support Incident Reporting Safety-Critical Systems***, www.dcs.gla.ac.uk/~johnson/papers/Safety_Science/software.html, 2001.

Biography

Nir Keren, Mary Kay O'Connor Process Safety Center, Chemical Engineering Department, The Texas A&M University System - 3574 TAMU; College Station, Texas 77843-3574, USA; Telephone - (979) 845-4950; Fax - (979) 458-1493; e-mail – mailto:nir@tamu.edu.

Nir Keren is pursuing a Ph.D. degree in the Interdisciplinary Engineering program at the Mary Kay O'Connor Process Safety Center, Texas A&M University, at College Station Texas. His primary interests include Industrial Safety Performance Measurements and novel applications of databases. Nir Keren completed his undergraduate studies at the Department of Mechanical Engineering at the Ben-Gurion University of the Negev, Israel. For 11 years, prior to joining the Center, Nir held positions as a Maintenance Manager, Project Manager, and Division Safety Engineer for several companies in Israel.

T. Michael O'Connor, Mary Kay O'Connor Process Safety Center, Chemical Engineering Department, The Texas A&M University System - 3122 TAMU; College Station, Texas 77843-3122, USA; Telephone - (979) 845-3489; Fax - (979) 845-6446; e-mail – mailto:tmo@che.tamu.edu.

Michael O'Connor is a Research Associate at the Mary Kay O'Connor Process Safety Center at

Texas A&M University. His research interests include high temperature heat exchangers and furnaces in ethylene and ammonia plants and metallurgy associated with these applications. Other areas of interest include incident database analyses and inherently safer design.

M. Sam Mannan, Mary Kay O'Connor Process Safety Center, Chemical Engineering Department, The Texas A&M University System - 3122 TAMU; College Station, Texas 77843-3122, USA; Telephone - (979) 862-3985; Fax - (979) 845-6446; e-mail – mailto:mannan@tamu.edu.

Dr. M. Sam Mannan is Professor of Chemical Engineering and Director of the Mary Kay O'Connor Process Safety Center at Texas A&M University. He is an internationally recognized expert on process safety and risk assessment. His research interests include hazard assessment and risk analysis, modeling of flammable and toxic gas cloud dispersion, inherently safer design, reactive chemicals and runaway reactions, aerosols, and abnormal situation management.