Safety Science 50 (2012) 1431–1437

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/ssci

Accident data for the Semantic Web

Ludwig Benner Jr. *

Starline Software Ltd., 12101 Toreador Lane, Oakton, VA 22124, USA

ARTICLE INFO

Article history:
Available online 21 January 2010

Keywords:
Lessons learned
Lessons learning system
Behavioral event sets
Accident investigation data structure
Event sets

ABSTRACT

This paper describes concerns about the documentation, dissemination and use of lessons learned from mishap investigations, impediments posed by current practices, and opportunities for improvement. Lessons are presently developed, documented and stored primarily in narrative form and relational databases, and disseminated in many forms and media, including the Internet. Current practices pose many impediments to maximized development, dissemination and use. Investigation process research and new data concepts behind the Semantic Web, exploited elsewhere, offer potential opportunities to overcome these impediments. To exploit these opportunities, formation of a working group to develop an improved Semantic Web-friendly mishap investigation lessons learning system is proposed. An example illustrating an alternative approach is described to support a reasonable expectation that an alternative lessons learning system could be developed.

1. Introduction: The need

The need to improve adaptive dynamic behavior of socio-technical systems through investigations of accidents, both before and after they happen, has long been of interest (Johnson, 1999). There is even a Society for Effective Lessons Learned Sharing (SELLS) (US Department of Energy, 2003). Maximizing development, dissemination and utilization of "lessons learned" is a continuing quest in many circles (US National Highway Traffic Safety Administration, 2003; Werner and Perry, 2004). One US report describes the need this way:

"NASA stated that it must do a better job of communicating the various lessons learned sources to employees, improving mechanisms to link these sources, and ensuring appropriate training for employees in order to maximize lessons learning"(United States General Accounting Office, 2002).

Some organizations have established lessons learned "centers" or operating feedback systems. They make use of mishap data inputs and inputs from other sources to generate databases with lessons learned for use within those organizations (US Army Combined Arms Center, Center for Army Lessons Learned, n.d.) or by recognized organizations and persons (National Advanced Fire and Resource Institute, 2007; European Commission, 2001). The lessons learned databases focus primarily on activities within the organizations' scope (Dien and Llroy, 2004).

Current investigation practices produce many kinds of outputs containing lessons learned. These outputs range from narrative reports, charts, completed forms, statistical trends or relationships, summary tables and books to bulletins, recommendation letters, check lists, training materials, or e-mail alerts. These outputs are derived by investigators or analysts who draw conclusions from the *investigation or incident data.*

Personal use of public or private lessons learned data is unknown or unreported, quantitatively, but interest in and use to generate new behaviors by individuals seems very limited. For example, one widely respected and emulated public incident lessons learned database with over 700,000 records had 88 search requests by individuals during a recent 6-year period (National Aeronautics and Space Administration, 2005). How many individuals in world process industries would buy a 334 page, $US80 book (Kletz, 2001) to find lessons learned that might apply to their tasks and then internalize all of them to change their behaviors? How frequently do individuals change their behaviors due to desired interpretations of generalized training, procedures, standards or regulations? Nobody knows. Data about acceptance of recommendations does not address whether lessons from investigations actually produced changed behaviors that improved safety, so assessments of present practices must rely on anecdotal evidence of users and observers or investigators. However, few would argue that present practices maximize investigation lessons learned dissemination and their use to bring about changes by all who might benefit from the data.

These circumstances suggest that prevailing lessons learning practices for the development, communication and use of lessons learned from mishap investigations merit examination. The examination should determine impediments to better performance, if a better lessons learning system might be developed, and how that might be accomplished.

* Tel.: +1 703 620 2270; mobile: +1 703 508 0847.
E-mail addresses: lbjrus@gmail.com, luben@mac.com

0925-7535/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ssci.2009.12.013

2. Present lessons learned practices

What is the present lessons learned "system" and why does not it maximize learning from current data?

2.1. Mishap lessons learned practices

Contemporary lessons learned practices reflect various "accident causation models." At present, investigators acquire, document and report "facts" or data in many forms and formats, in many diverse and often isolated systems (Sklet, 2003). These data are used by investigators and analysts to piece together a description and explanation of what happened, usually in narratives or on forms, using natural language. Such accident data also form the basis for conclusions about causes, cause factors, root causes, and other cause-oriented findings, from which investigators and analysts derive findings and recommendations. Findings and recommendations constitute the "lessons learned" from an investigation. Analysts then abstract, code, characterize, aggregate or otherwise refine or condense them. They are then "published" internally or made public in various kinds of media as reports, articles, papers, books, stories, graphics, training materials, check lists, etc. They also find their way into procedures or standards and regulations. The "published" data are then preserved by storage in organizational files or computerized databases for retrieval and subsequent uses at a later date.

Dissemination practices vary, but generally can be categorized as (a) electronic and (b) non-electronic written, verbal and graphic dissemination. Electronic dissemination is achieved with computers and computerized databases, e-mails, and Internet sites. Non-electronic dissemination is achieved through published or internal investigation reports, tables, checklists, on-the-job training, safety meetings, standards, training sessions, codes or regulations, and books, for example. End users' learning and ultimate changes depend on the content, access to and assessment of these lessons plus other considerations and tradeoffs, but they would not occur at all without availability of the investigation lessons.

Investigation data are also used for research to develop lessons learned in the form of historical trends or statistical correlations, using statistical analyses or data mining techniques. The data are also frequently abstracted or characterized to generate lists of causes and causal factors referenced in investigation report databases, safety digests and investigation software.

2.2. Impediments to learning

What are the shortcomings of the present lessons learned practices?

A 2004 paper (Werner and Perry, 2004) cited several observed barriers to effective *capture and use* of investigation lessons learned. These barriers could be summarized as:

- Lessons are not routinely identified, collected and shared across organizations and industries;
- Unorganized lessons are too difficult to use, because there is too much material to search, it may be formatted differently for different reports, it is not quickly available, or work pressures do not allow time or resources to find it;
- Reuse is rather ad hoc and unplanned;
- It is often hard to know what to search for or how to find useful documents; and
- Taking time to search for, identify, access and then learn from them within an organization is a problem.

Users and managers identified additional impediments, including irrelevance, cycling of a company practice or instruction, repetitive lessons, suspect tools, and lack of evidence that lessons are being applied toward future success (Cowles, 2004).

No previous substantive research addressing the *development of lessons learned* during investigations is known to exist. Analyses and criticisms of contemporary investigation practices abound in the literature.[1] Investigation problems such as investigative perspectives, conflicting objectives, flawed assumptions, scope, biased data selection, interpretation or representation of observations, logic errors, vocabulary, language ambiguities, premature conclusions, quality control, recommendation development and implementation, and overlooked lessons learned problems have been reported in detail (Hendrick and Benner, 1986). Each investigation problem contributes to flawed development and use of lessons found during investigations.

Personal observations during investigations over a 35-year span, impediments cited above and analysis of reasons investigation recommendations were not implemented, suggest several underlying impediments preventing maximized development and dissemination of lessons learned. These underlying impediments could be characterized as:

- Current perceptions of investigation data needs that limit data presently available for sharing;
- Natural language barriers that lead to diverse source data content and structures, impeding identification of relevant behaviors as lessons;
- Data that are lost due to software obsolescence; and
- Liability concerns that motivate a desire to withhold accident data from publicly accessible sources.

2.2.1. Perceptions of data needs

Perceptions of what investigation data should be acquired and disseminated are based on contemporary "accident causation models." These models and the view of the accident phenomenon behind them may be the greatest impediment to learning. Investigation models, purposes or mandates shape those perceptions. Investigation processes are not designed with the goal of informing all those who need to initiate new behaviors. Currently investigation practices focus on determining the cause or cause factors, multiple causes, problems, and "root" causes, for example, from which investigators or analysts infer lessons learned to address with recommendations. Outputs typically do not provide lessons data in a form from which individuals can quickly derive the specific behavioral changes they need to make. In other words, the target audience is spoon-fed selected changes deemed desirable by the "experts," in the form of recommendations.

2.2.2. Natural language barriers

The preponderance of current accident data is documented using natural language, rather than a "professional language" like those that exist in mathematics or music or medicine or other professional fields. This usage tolerates wide variations in the vocabulary, morphology, syntax, meaning, context and level of abstraction of documented investigation data. That variability impedes manual analysis, machine comparisons and tabulations or rule-based manipulation such as rational concatenation of elements, or interoperability, machine access and machine presentations of the data.

In these circumstances, many investigation data schemes provide accident data definitions, to indicate intent and improve consistency. Data improvement efforts have typically been directed at enhancing data uniformity of meaning, with guides, dictionaries, glossaries or check lists defining words and terms (European Community, 2006). However, most lack a defined data *structure*

[1] Many such papers are found at http://www.iprr.org.