CHEMICAL ENGINEERING TRANSACTIONS
VOL. 116, 2025
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.cetjournal.it
Guest Editors: Bruno Fabiano, Valerio Cozzani

ISBN 979-12-81206-18-2; **ISSN** 2283-9216
DOI: 10.3303/CET25116011

Using Accident Anatomy Analysis and Small Language Models to Support Systematic Lessons Learned Analysis and Improve the Completeness of Hazard Analysis

J.Robert Taylor

The Centre of Maritime Health and Society (CMSS), University of Southern Denmark
roberttayloritsa@gmail.com

Over a period of 36 years, risk analysis was made of 105 process plants using HAZOP, action error analysis, mechanical integrity audit or pre-commissioning audit and QRA. Follow up studies were made at the end of the period. Altogether 82 major hazards accidents occurred, with a total of 168 fatalities as consequence. Of these 50% were predicted but not prevented due to inadequate safety management. The other 50% were due to lack of knowledge of some rare hazard types at the time of risk assessment.
To improve completeness of hazard identification the Systematic Lessons Learned method was used which consists of cross referencing a large database of accident case histories. To ease the preparation of the database, a program, Cassandra, was developed for semantic analysis of accident case histories and to allow natural language retrieval of supplementary information for hazard identification
The risk analyses were made using HAZOP with the workshop teams including the plant or design engineers and senior operators, with the author as facilitator. Additionally, generic action error analyses for operation of process unit types to support human error analysis (Taylor 2016). Completeness of analyses was checked using automated HAZOP analysis (Taylor 2017) and automated cross checking between manually and automatically completed results.

1. Introduction

Over a period of 36 years, 105 risk analyses were made for oil, gas and chemical plants including several refineries and three petrochemical complexes. In all the studies covered 449 process units and 7134 process unit years of operation as of ultimo 2014. Thes plants were situated in a wide range of countries, Alaska, California, Texas, Venezuela, Brazil, the Middle East, North Africa, Mainland Europe and Scandinavia. Several complete refineries, three petrochemical complexes and several large pharmaceutical plants are included in the dtatse. Recommendations for risk reduction were made in all cases.
In the plants which had been analysed, 82 major hazards type accidents occurred, defined here as process accidents with consequences for persons outside the process unit boundary. Of these, 41 were predicted but not prevented due to lack of implementation of safety recommendations. These are described in (Taylor 2018). For just one predicted accident, the risk was originally assessed as ALARP. 38 of the accidents were not predictable at the time of analysis due to lack of knowledge.
192 major risk reduction recommendations were made prior to 1994. In 14 cases recommendations were either rejected or were not implemented for various reasons and accidents occurred which could have been prevented. This severity of this problem was recognised after two very severe accidents in 1994. The weakness of traditional risk analysis approaches in preventing these accidents was apparent. The presentation of risk analysis results was modified, with stronger presentations of consequences, illustrated with photographs of the actual plant situation and of earlier accidents, with possible conceptual designs for risk reduction and with full risk cost benefit analyses for the recommended risk reduction, with guidance on loss prevention techniques and generally with a conceptual design for risk reduction of the measures. As a result, 1536 recommendations were made and implemented between 1994 and 2016, with just 6 remaining unimplemented. Reasons for failure to implement recommendations since 1994 are given in figure 2.

Paper Received: 8 December 2024; Revised: 22 February 2025; Accepted: 20 April 2025
Please cite this article as: Taylor J., 2025, Using Accident Anatomy Analysis and Small Language Models to Support Systematic Lessons Learned Analysis and Improve the Completeness of Hazard Analysis, Chemical Engineering Transactions, 116, 61-66 DOI:10.3303/CET25116011

For plants with risk analyses prior to 1994 the major hazards accident (MHA) frequency was $1.3*10^{-2}$ per unit year. For the risk analyses completed after 1994 the number of MAH accidents was also $1.3*10^{-2}$ per unit year. Including the accidents prior to 1994 the largest death toll in one accident was 64 fatalities and the total number of fatalities was 146. The MAH accident frequency as observed for the full period was $1.15*10^{-2}$ per process unit year.
For plants with risk analyses completed before 1994, the total number of fatalities up to 2014 was 152, or 88 if one very severe pipeline accident is excluded. The fatality risk for these plant units was $2.3*10^{-2}$ per unit year, or $1.3*10^{-2}$ per unit year if the pipeline accident is excluded.
For plants with risk analyses completed after 1994 (i.e. with improved communication and improved loss prevention support) the number of fatalities was 16. Five of these were due to unpredictable accidents. The fatal accident rate for these plants was $4.8*10^{-3}$ per unit year.
Figure 1. shows the breakdown of the major hazard accident scenarios according to hazard analysis problem types.

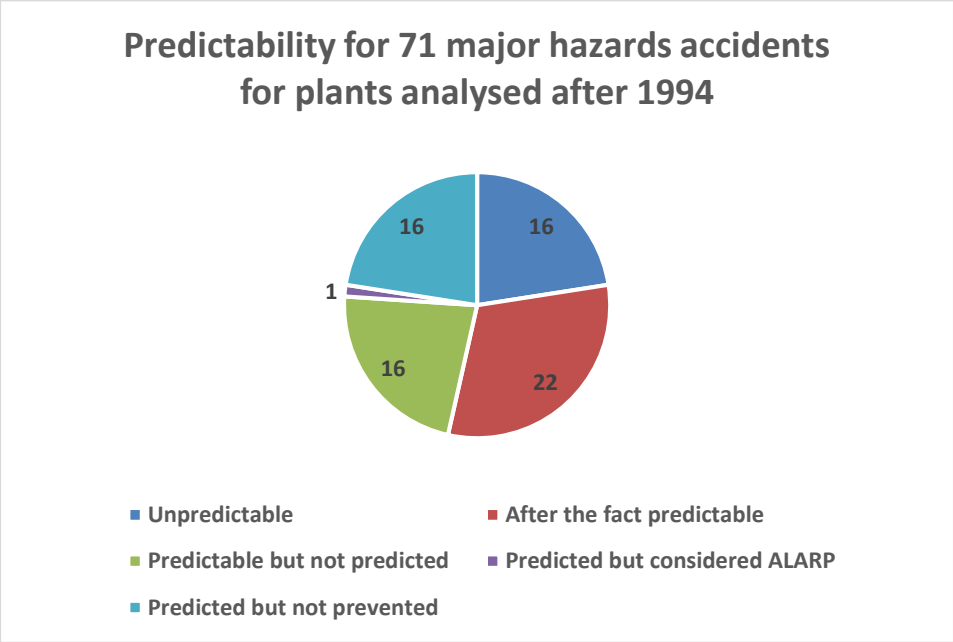

Figure 1 Predictable and unpredictable accidents occurring after 1994

The “after the fact predictable accidents” in figure 1 are those which can be predicted now on the basis of knowledge gained from the accident investigation. Predicted but not prevented accidents were not prevented for a variety of reasons, including management rejection of recommendations in one case, insufficient time for implementation of preventive measures, recommendations “lost in committee”, recommendations delayed “until next turn-round” and then overlooked etc. The predictable but not predicted accidents were those in which the scope of analysis was limited or the methods inadequate. The “predictable but not predicted” arose from limitations in the scope of risk analyses imposed on the studies by the client companies, mostly lack of human error analysis; or by limitations in methods used. Predictability was determined by means of follow-up automated analyses for all of the 105 plants carried out in 2015 (see Taylor 2017).
Unpredictable accidents are those which involve complicated causal mechanisms, accidents which still cannot be explained even after extensive investigation including laboratory analyses, and those which involve mechanisms such as crevice corrosion where causes can only be determined by dismantling equipment.
Methods used to predict the unpredicted accidents in the dataset in follow up by means of automated HAZOP studies are shown in figure 3. The efficacy of the methods was demonstrated by automated HAZOP analysis and automated design review during follow-up studies in 2014 to 2017..

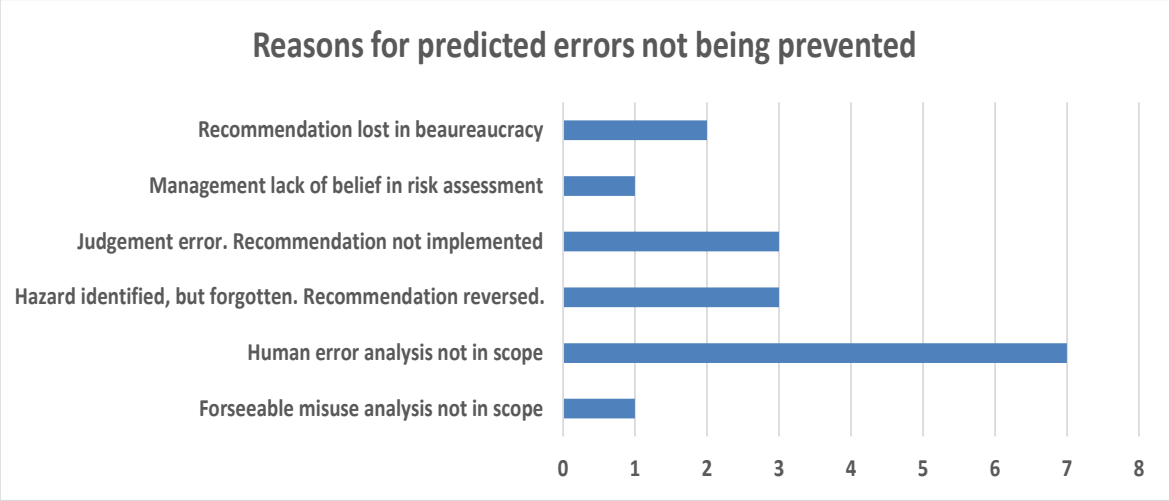

Figure2 Reasons why predicted accidents were not prevented in 16 MAH accidents

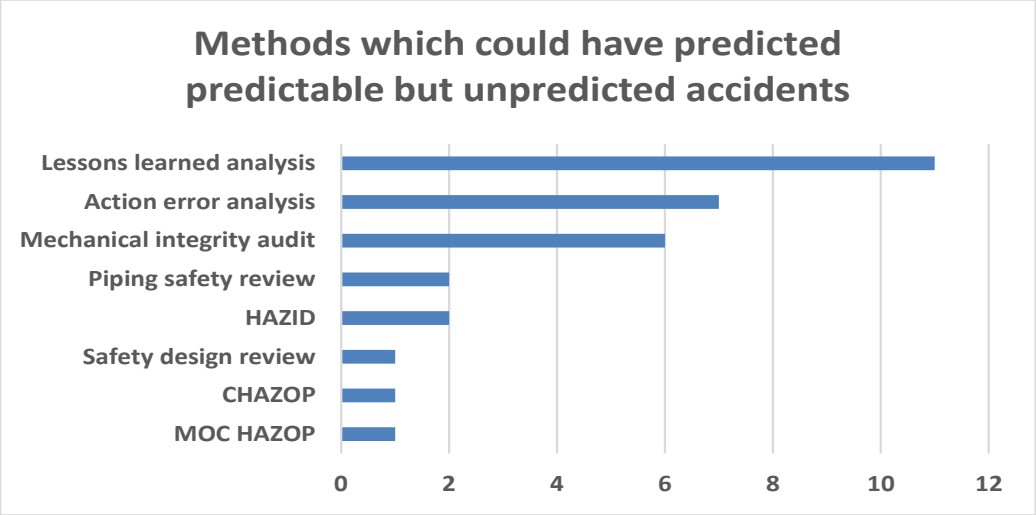

Figure 3 Methods which could have predicted the unpredicted but predictable accidents

The accidents occurring also show the importance of careful management of the hazard identification and risk reduction processes (Taylor 2020). Managements need to recognise the importance of a complete risk reduction chain and that a simple reliance on a QRA or HAZOP which is simply read, approved and then archived is not adequate to satisfy the duty of care requirement. Risk analyses are a useful tool which can be used to support continuing efforts in safety management, and reports should therefor be formulated to be readable, understandable, teachable for new employees and updatable as plants develop.

2. Completeness of hazard identification

An ideal measure of completeness for hazards identification would be the ratio of hazards identified to possible hazards. This is not a feasible measure however because the total number of hazards including future accidents cannot be known. A practical measure (Taylor 2012) is:

Historical completeness = No. hazards identified / No. hazards in a large case history collection

This completeness measure will change as more accident types are included in the data set, but in practice it has proved quite stable over the years. Completeness needs to be evaluated for hazard types which are actually identifiable from source material available. For example, HAZOP can achieve high degrees of completeness for

hazards deducible from piping and instrumentation diagrams but is not able to predict many piping hazard types which require piping drawings or 3D models for assessment.
Automated HAZOP (Taylor 2017), automated QRA and automated Action Error Analysis were used for determining quality of hazard identification and risk analysis, allowing for continued updating of the completeness and QRA results. HAZOP completeness was measured as between 85 and 95% for HAZPS completed in the 1980's, but with modern practice HAZOP can achieve completeness levels of 95 to 98% with a well facilitated team of experienced design and operations engineers but of course only for hazards which are visible from P&IDs, but can be improved to 99.8% by follow up with Lessons learned analysis as described below. Remember though that many accidents cannot be observed by studies of P&IDs alone.

3. Systematic Lessons learned analysis and Accident Anatomy Analysis

Systematic Lessons learned analysis is a method which uses a well indexed collection of accident case histories. It is used to supplement HAZOPs by answering question like "what have we missed?". In a collection by the author (Taylor 2015) the collection is indexed by equipment type, parameter deviation type, consequence type and materials involved. Systematic lessons learned analysis has allowed completeness of automatic HAZOPs to be increased to 98.8% as determined from follow up studies of all the 105 studies with the latest data collection in 2015.
Lessons learned records take time to read, may not be read, or may not be remembered even if read. Experiments during HAZOP studies showed that most people do not remember old accidents if they occurred more than a few years earlier (Kletz 1993). A visual approach provides better recall. The accident anatomy diagrams (see figure 4) are easy to remember by topic, and the details are readily obvious with little reading when they are retrieved.
There as many sources of good accident case histories including the US Chemical Safety Board, the UK HSE, the French ARIA database and the Japanese Knowledge Base. Such case histories can be indexed and classified. However the classification studies can represent only a small fraction of the valuable knowledge in the accident reports. The Accident Anatomy method was developed at Risø National Laboratory (Bruun, Rasmussen and Taylor 1979) to make better use of accident reports information. It involves the following steps:

1. Select a set of accident reports which have a degree of similarity and the same key event.
2. For each accident, draw a cause consequence diagram, aiming for consistency of terms.
3. Combine the diagrams so that one overall picture of accident possibilities is obtained
4. For each path through the diagram the number of cases which follow the path.
5. Redraw the diagram with thickened connecting lines depending on the number of cases following each path.
6. Mark up case references on the diagram.
7. Add safety rules of thumb and regulations which apply for each stage of the path.
8. Index the diagrams for easy retrieval

Figure 4 shows an accident anatomy diagram for runaway reactions in a kettle reactor. The diagram shows a quite large overlap with the statistical data collection by (Barton and Nolan 1989)

4. Cassandra – a small language model for accident anatomy diagram creation and referencing

Lessons learned analysis and accident anatomy diagrams are good tools for improving hazard identification. However they are very expensive to produce in terms of man-hours. Faced with several thousand case histories to process, an automated tool was developed for analysing and summarising accident reports. Large language models were considered for the task, but these have drawbacks. The suffer from "hallucinations" and they provide statistical correlations, not understanding (Qi et al.2023). Also, they require massive online resources and much time to train.
Instead, an approach which we call "small language modelling" was used. This uses old techniques from the 1960's and 70's, augmented transition networks (ATNs) for syntax analysis (Winograd 1983, Allen 1987). It uses system theoretic models for the knowledge base, an approach which has proved effective in over 400 automated HAZOP studies. The methods used have been outdated for language translation purposes, being superseded first by statistical methods and more recently by generative neural networks. However with the improvements in rule based techniques over 40 years, the use of semantic networks based on systems theory, they can be very efficient for technical purposes.

The program, Cassandra, uses a subset of the English language called “Engineering English”. It uses a vocabulary of about 50,000 words and can store knowledge in the form of up to 1,000,000 sentences (an arbitrary limit), sufficient for about 500 books of 500 pages each. As can be seen, the language model is “small” only in the sense that its resources are small (a laptop with 8Gb of main memory) and in its very low energy use, an issue which is becoming important for large language models.

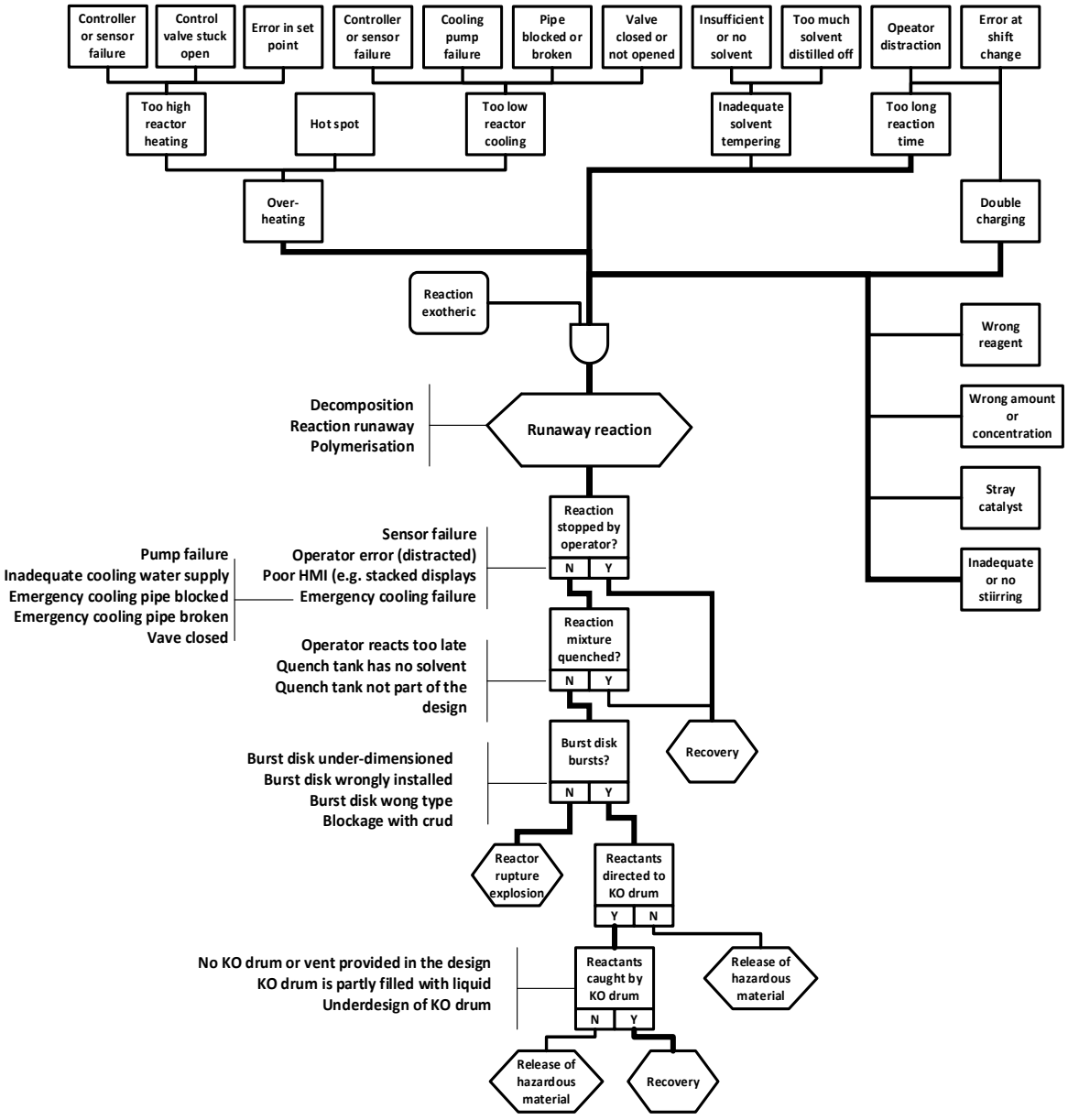

Figure 4 Accident Anatomy diagram for 62 runaway reaction accidents (frequent paths emphasised)

6. Results and future work

The diagram in figure 4 was constructed automatically from 62 accident report texts, varying in length from between 1/3 of a page to a full page. The processing took just under 2 seconds and required about $3*10^{-4}$ kW hours of energy, about 0,01% of that required by a modern large language model. Why is there such a high increase in efficiency compared with an LLM? This appears to be the result of the use of system theoretic

semantics, which allows for an efficient mechanistic knowledge base. The context on which question answers are based is the entire corpus of knowledge in the knowledge base.
The software is completely dependent on the source texts and cannot “invent” new accident types, as current large language models can. However it can show novel accident event scenarios by combining known facts using physical principles.
The value of accident anatomy studies has been demonstrated in a series of reports in different areas of engineering and has become a standard tool for safety studies. Other applications of the techniques are planned however, including checking of design specifications.
One of the most difficult areas in interpreting accident reports is understanding, categorising and retrieving the parts which describe the human aspects of accident, including human cases and failure of human emergency response. The original implementation of Cassandra could only create knowledge bases derived from system theoretic models i.e. mechanistic information. The newest developments make use of human cognitive models to describe communication, understanding, decision making, and allocation of attention etc. to allow the “human parts” of accident reports to be included in the Cassandra knowledge base.

References

Allen J. (1987) Natural Language Understanding, The Benjamin/Cummings Publishing Company
Barton J.A. and P F Nolan (1989) Incidents in the Chemical Industry Due To Thermal-Runaway Chemical Reactions, IChemE SYMPOSIUM SERIES No. 115
Bruun, O., Rasmussen, A., & Taylor, J. R. (1979). Cause consequence reporting for accident reduction. The accident anatomy method. Risø-M No. 2206
Kletz, T.A. (1993) Lessons from Disaster: How Organizations Have No Memory and Accidents Recur, Gulf Publishing Company
Qi Y., Xingyu Zhao, Diddartha Khastgir, Xiaowei Huang (2023) Safety Analysis in the Era of Large Language Models, https://doi.org/10.48550/arXiv.2304.01246
Taylor J.R. (2012) Forty Years of HAZOP, Loss Prevention Bulletin, October 2012
Taylor J.R. (2015) Systematic Lessons Learned Analysis for Oil and Gas Plant, QRA Quality Report Vol 26, //www.academia.edu/35376594/Systematic_Lessons_Learned_Analysis_for_Oil_and_gas_Plant
Taylor J. R, (2016) Human Error in Process Plant Design and Operation, CRC Press
Taylor J. R. (2017) Automated HAZOP Revisited, Process Safety and Environmental Protection 111 635–651
Taylor J.R. (2020) Organisational Failure Analysis for Industrial Safety, ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering, March 2020, Vol. 6 / 011006-3
Winograd T. (1983) Language as a Cognitive Process, Addison Wesley

Sources of accident data

US Chemical Safety Board, https://www.csb.gov, (126 reports on line)
The ARIA Database - La référence du retour d'expérience sur developpement-durable.gouv.fr
https://www.aria.developpement-durable.gouv.fr (204 accident records online)
Failure Knowledge Database, Japan & Science Technology Agency, Japan. http://shippai.jst.go.jp/en/Search. New host http://www.shippai.org/fkd/en/lisen/cat102.html (142 chemical industry records online)
Kletz T (2009) What went Wrong, 5th edition, Elsevier (90 topics, many with several examples)
Taylor J.R. (2015) Systematic Lessons Learned Analysis for Oil and Gas Plant, QRA Quality Report Vol 26, //www.academia.edu/35376594/Systematic_Lessons_Learned_Analysis_for_Oil_and_gas_Plant, (139 accidents, 426 lessons learned)

.