Proceedings of the 32nd European Safety and Reliability Conference (ESREL 2022)
Edited by Maria Chiara Leva, Edoardo Patelli, Luca Podofillini, and Simon Wilson

Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-5183-4_S29-05-616-cd

A Taxonomy for Modelling Reports of Process Safety Events in the Oil and Gas Industry

Dario Valcamonico

Department of Energy, Politecnico di Milano, Milan, Italy. E-mail: dario.valcamonico@polimi.it

Piero Baraldi

Department of Energy, Politecnico di Milano, Milan, Italy. E-mail: dario.valcamonico@polimi.it

Enrico Zio

Department of Energy, Politecnico di Milano, Italy. E-mail: enrico.zio@polimi.it
MINES Paris-PSL, Centre de Recherche sur les Risques et les Crises (CRC), Sophia Antipolis, France

Anna Crivellari

Eni NR division, Milan, Italy. E-mail: anna.crivellari@eni.it

Luca Decarli

Eni NR division, Milan, Italy. E-mail: luca.decarli@eni.it

Laura La Rosa

Eni NR division, Milan, Italy. E-mail: laura.larosa@eni.it

Abstract:
The Process Safety Management System (PSMS) of an industrial asset relies on multiple and independent barriers for preventing the occurrence of major accidents and/or mitigating their consequences on people, environment, asset and company reputation. It is, then, fundamental to assess the performance of the barriers with respect to the occurrence of Process Safety Events (PSEs), i.e. unplanned or uncontrolled events during which a Loss Of Primary Containment (LOPC) of any material, including non-toxic and non-flammable material, occurs. An essential aspect of PSMS is learning from incidents and taking corrective actions to prevent their recurrence. For this, a procedure for timely and consistently reporting and investigating PSEs is generally implemented. After the occurrence of a PSE, a report containing free-text and multiple-choice fields is filed to describe the PSE, its causes and consequences, and to provide a quantification of its the level of severity with reference to predefined Tier levels, as per API RP 754 guidelines. This work investigates the possibility of text-mining and structuring the knowledge on the performance of the PSMS from an electronic repository of PSE reports. The methodology developed falls within the framework of Natural Language Processing (NLP), combining Term Frequency Inverse Document Frequency (TFIDF) and Normalized Pointwise Mutual Information (NPMI) for the automatic extraction of keywords from the PSE reports. Then, a taxonomy is built to organize the vocabulary in a top-down structure of homogeneous categories, such that semantic and functional relations between and within them can be defined. Based on these relations, a Bayesian Network (BN) is developed for modeling the PSEs consequences. The proposed methodology is applied to a repository of real reports concerning the PSEs of hydrocarbon facilities of an Oil and Gas (O&G) company.

Keywords: Process Safety Events, Natural Language Processing, Text mining, Taxonomy, Bayesian Network.

1. Introduction

Safety of people, environment and asset from major accidents in hydrocarbon facilities is achieved with preventive and mitigative barriers. The assessment of the performance of the barriers is, then (Decarli et al. 2020).

Traditionally Risk Assessments (RAs) are based on hazard identification, evaluation of accident probabilities and consequences, and risk estimation (Feng et al. 2012; Vinnem et al. 2012; Zio 2018). These steps are typically tackled with a systematic analysis of the available information performed by multidisciplinary teams of experts with techniques such as Bow-Tie (Khakzad, Khan, and Amyotte 2013), Dynamic Risk Assessment based on statistical failure data (Zeng and Zio 2018) and Bayesian Networks (BN) (Li, Chen, and Zhu 2016).

Taking advantage of recent developments in the field of Artificial Intelligence (AI), this work considers the application of AI to an electronic repository of textual reports of Process Safety Events (PSEs), i.e. unplanned or uncontrolled events during which a Loss Of Primary Containment (LOPC) of any substance, including non-toxic and non-flammable material, occurs ("API RP 754" 2021).

These reports describe the PSEs, their causes and consequences, and quantify the levels of severities with reference to predefined Tier levels ("API RP 754" 2021). They are extended investigation reports originated from flash reports issued after the occurrence of the event, and are compiled by system operators in free text and multiple-choice fields.

The objective of the work is to systematically mine out the information that the textual reports contain about the PSE root causes and the lessons learned from them, to model PSEs and estimate their probabilities (Sattari et al. 2021).

The developed methodology is based on the combination of a Natural Language Processing (NLP) technique for the identification of keywords describing the factors influencing the occurrence and severity of PSEs, and a BN (Jensen and Nielsen 2007) for modeling the keywords, the PSE types, causes and consequences. The identified keywords are systematically assigned to the barriers by developing a taxonomy of the vocabulary used in the reports.

The proposed methodology has been applied to a real repository containing multilingual reports of PSEs (in the order of thousands) occurred in hydrocarbon facilities of an Oil and Gas (O&G) company.

The remaining of this paper is organized as follows: Section 2 illustrates the problem statement; Section 3 describes the proposed methodology; Sections 4 introduces the case study; Section 5 discusses the obtained results; in Section 6, conclusions and remarks are drawn.

2. Problem Statement

Hydrocarbon industry typically collects reports of PSEs in structured electronic databases containing free texts and multiple-choice fields (Milana et al. 2019). Typically, the incidents databases are organized in a structured way, where the generic i^{th} report, r_i, is recorded in the form of: *a*) a description in free text, d_i, of the PSE; *b*) a multiple-choice entry, $x_i \in \{x^1, \dots, x^A\}$, which indicates the PSE cause among a predefined set of A alternatives; *c*) a multiple-choice entry, $y_i \in \{y^1, \dots, y^B\}$, which indicates the type of event among a predefined set of B alternatives; d) a multiple-choice entry, $z_i \in \{z^1, \dots, z^C\}$, which indicates the Tier level, i.e. an indicator of the severity of the consequence of the PSE in a scale with C levels. We assume that a repository of D reports of PSEs, $R = \{r_i, i = 1, \dots, D\}$, is available.

The objective of this work is the development of a methodology to exploit the information contained in the repository R. Specifically, we aim at building a probabilistic model of the PSE consequences (z) as a function of its causes (x), type (y) and states of the barriers (w). Once the model has been developed, it can be used to support risk analysis for the estimation of the consequences of a PSE of a given cause x, a given type y, or PSEs with different combinations of the barrier states (w).

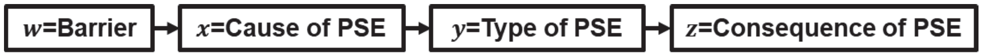

Figure 1: Example of a logic model of PSEs, constituted by the elements of cause, type, consequence, and influencing factors of the PSE.

3. Methodology

The causal model of a PSE shown in Figure 1, which links the states of the barriers, w, the causes of the PSE, x, the types of PSE, y, and the PSE consequences z, is transformed into a BN. This choice is motivated by the fact that BNs are probabilistic models in which the causal relations among the variables are represented in terms of conditional probabilities (Di Maio et al. 2021). Section 3.1 describes the architecture of the BN, Section 3.2 the method developed to estimate its Conditional Probabilities Tables (CPTs) and Section 3.3 its use.

3.1 BN architecture

The architecture of the BN is obtained combining the causal model of the PSE with the categorical and textual data of the repository R. Specifically, the nodes are the barriers w, the cause of the PSE, x, the type of PSE, y, and the consequence of the PSE, z. The states of the nodes x, y and z are the multiple-choice entries of the repository R, i.e. $\{x^1, \dots, x^A\}$, $\{y^1, \dots, y^B\}$ and $\{z^1, \dots, z^C\}$, respectively. With respect to the definition of the nodes referring to the barriers, a procedure based on the definition of a taxonomy of the barriers is developed. The idea behind the use of the taxonomy is to define a set of M states $\{w^1, \dots, w^m, \dots, w^M\}$ and to assign the words of the vocabulary of the PSE reports $\{r_i, i = 1, \dots, D\}$ to a specific state w^m. Sections 3.1.1 describes the taxonomy, Section 3.1.2 the method to extract the vocabulary from the reports and Section 3.1.3 the procedure followed to associate the words of the vocabulary to the taxonomy. Figure 2 shows the obtained BN and the sources of information used to define it.

3.1.1 Taxonomy of the barriers

The process of abstraction and organization of concepts is recognized to be the starting point for organically structuring the domain knowledge in many fields, such as safety in healthcare (Itoh, Omata, and Andersen 2009), biotechnology (Parks et al. 2018) and economic sustainability (Saidani et al. 2019). Taxonomies have been developed for this scope, in applications involving systems management (Mackie, Welsh, and Lee 2006), system aging (Ansaldi et al. 2020), system safety (Hodkiewicz et al. 2021) and system design (Ahmed, Kim, and Wallace 2005). In general, a taxonomy is a top-down hierarchical structure based on sub-levels to organize the knowledge in a domain. Each sub-level represents a group of entities that share common properties. For example, in the biology domain, traditional taxonomies for biological organisms organize species based on hierarchically organized morphological characteristics (number of legs, shape of head, etc.) (Dayrat 2005). In the context of the present work, we develop a taxonomy to systematically structure the language used by the operators to describe the barriers of the system in the reports. It is based on two sub-levels: *i*) the possible types of barriers $w^1, \dots, w^m, \dots, w^M$, with M indicating the number of types of barriers; *ii*) the words used in the repository to refer to the barriers, e.g., $v_1^m, \dots, v_j^m, \dots, v_{V^m}^m$ for a generic barrier of type w^m, with V^m indicating the number of words of the vocabulary referring to a barrier of type w^m. Specifically, the barriers are organized in types (sublevel *i*) by system experts according to their knowledge. The definition of sub-levels *ii*) requires the identification of the vocabulary of the repository (Section 3.1.2) and the assignment of each token of the vocabulary to a barrier type (Section 3.1.3). Figure 3 shows a graphical representation of the taxonomy of the barriers.

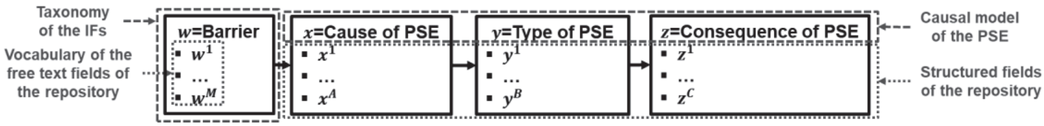

Figure 2: Obtained BN and sources of information used to define it.

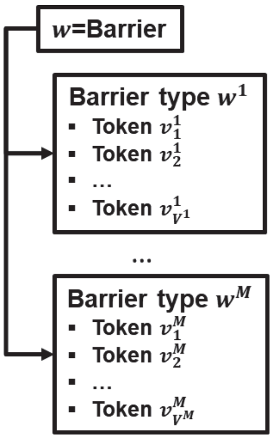

Figure 3: Conceptualization of the taxonomy of the Barrier.

3.1.2 Vocabulary

The vocabulary of the repository R is the set of unique words and their contiguous combinations, $\{v_j, j = 1, \dots, V\}$, used in the reports. It is obtained by converting the free text description of the PSE of each report, d_i, $i = 1, \dots, D$, into the corresponding set of tokens $\tilde{d}_i$, $i = 1, \dots, D$, following the procedure described in (Valcamonico, Baraldi, and Zio 2021), which is based on the steps of tokenization, cleaning of the text, lemmatization and identification of commonly used n-grams, i.e. sequences of contiguous words which allow improving the semantic clarity and the precision of the language (e.g. "pressure_safety_valve" and "lack_of_maintenance"). The method here used for the identification of n-grams is based on the use of the Normalized Pointwise Mutual Information (NPMI) index, which measures the amount of information that a specific combination of words carry with respect to the information of the words alone (Bouma 2009). In practice, a n-gram is included in the vocabulary only if its associated NPMI index is larger than a threshold $thresh_{NPMI}$, which is the only hyperparameter of the method used for the definition of the vocabulary.

3.1.3 Definition of the types of barriers and identification of their corresponding tokens

The definition of the types of barrier $\{w^1, \dots, w^m, \dots, w^M\}$ is performed by field experts through a process where the tokens $\{v_j, j = 1, \dots, V\}$ of the vocabulary are grouped considering the similarity of their semantic meaning and, then, assigned to the proper type of barrier w^m.

3.2 CPTs estimation

CPTs of the BNs are typically estimated by combining expert knowledge with data (Xiaoguang, Yu, and Zhigao 2019). In this work, which is devoted to the exploitation of the textual data, we consider the estimation of the CPTs using only the information in the repository R of the PSEs. Future work will investigate the possibility of integrating this source of information with expert knowledge or other RA models.

3.2.1 Identification of the keywords

The objective is the extraction of the set of N keywords, v_k, $k = 1, \dots, N$, among the V tokens of the vocabulary $\{v_j, j = 1, \dots, V\}$, with $N \leq V$. To this purpose, a method based on Term Frequency Inverse Document Frequency (TFIDF) has been developed in (Valcamonico, Baraldi, and Zio 2021). The main idea is to use TFIDF to transform the set of tokens in each report $\tilde{d}_i$, $i = 1, \dots, D$, into numerical vectors h^i, $i = 1, \dots, D$, where the generic element h_j^i is associated to a measure of semantic importance of the token v_j in report d_i. Then, the keywords extracted from report d_i are the tokens that satisfy $h_j^i \geq thresh_h$, where the method hyperparameter $thresh_h$ indicates the minimum value of semantic importance needed to obtain a keyword. Finally, the total set of keywords $\{v_k, k = 1, \dots, N\}$ is made by the tokens which have been selected for at least one report d_i, $i = 1, \dots, D$. Notice that, since the tokens of the vocabulary are organized in the taxonomy, the types of barriers $\{w^1, \dots, w^m, \dots, w^M\}$ associated to the keywords are automatically obtained by using the taxonomy.

3.2.2 Estimation of the conditional probabilities of the CPTs

The CPTs of the BN are estimated using the information in the repository R. The conditional probability elements $P(y^b|x^a)$ and $P(z^c|y^b)$ of the CPTs between nodes x and y, and between nodes y and z, respectively, are estimated by (Rohmer 2020):

$$P(y^b|x^a) = \frac{\rho(x^a, y^b)}{\rho(x^a)} \tag{1}$$

$$P(z^c|y^b) = \frac{\rho(y^b, z^c)}{\rho(y^b)} \tag{2}$$

where $\rho(x^a, y^b)$ ($\rho(y^b, z^c)$) is the number of reports in which the node y (z) is in the state y^b (z^c), and the immediate parent nodes x (y) is in state x^a (y^b), and $\rho(x^a)$ ($\rho(y^b)$) is the number of reports in which the immediate parent nodes x (y) is in state x^a (y^b). Considering the CPT, for the node w and the node x, the conditional element $P(x^a|w^m)$ is estimated by:

$$P(x^a|w^m) = \frac{\rho(x^a, w^m)}{\rho(w^m)} \tag{3}$$

where $\rho(x^a, w^m)$ is the number of reports associated to x^a and containing at least one keyword associated to the type of barrier w^m and $\rho(w^m)$ is the number of reports containing at least one keyword assigned to w^m.

3.3 Application of the BN to RA

The developed BN model can be used to compute quantities of interest in RA. Some useful probabilities that can be estimated are:

(i) the unconditional probability of occurrence of a PSE with class of severity $z^c, c = 1, \dots, C$:

$$P(z^c) = \sum_{b=1}^{B} P(z^c|y^b)P(y^b) \tag{4}$$

where $P(y^b) = \sum_{a=1}^{A} P(y^b|x^a)P(x^a)$ and $P(x^a) = \sum_{m=1}^{M} P(x^a|w^m)P(w^m)$.

(ii) the probability of occurrence of a PSE with class of severity z^c, $c = 1, \dots, C$, for a PSE of type y^b, $b = 1, \dots, B$. This is obtained by estimating the probability of the consequences z^c of a PSE conditional to the type of event y^b, i.e. $P(z^c|y^b)$, which coincides with the element corresponding to state z^c of node z and state y^b of node y of the CPT from node y to node z.

(iii) the probability that a PSE of type y^b, $b = 1, \dots, B$, has been caused by x^a, $a = 1, \dots, A$. This is obtained by estimating the probability of the cause x^a, conditional to a PSE of type y^b:

$$P(x^a|y^b) = \frac{P(y^b|x^a)P(x^a)}{\sum_{a=1}^{A} P(y^b|x^a)\, P(x^a)} \tag{5}$$

where $P(x^a) = \sum_{m=1}^{M} P(x^a|w^m)\, P(w^m)$.

(iv) the probability that the barrier w is of type $w^m, m = 1, \dots, M$, given that the cause of the PSE is x^a, $a = 1, \dots, A$. This is obtained by estimating the probability of the type w^m conditional to the cause x^a:

$$P(w^m|x^a) = \frac{P(x^a|w^m)P(w^m)}{\sum_{m=1}^{M} P(x^a|w^m)P(w^m)} \tag{6}$$

4. Case Study

In an electronic repository D of reports of PSEs occurred in hydrocarbon facilities, each report is composed by:

(i) a free-text, d, written in English by a system operator and containing the description of the event and of the factors influencing the PSE;

(ii) the cause, x, which can be x^1="Human", if the PSE was caused by system operators incorrectly following procedures or misusing the available and prepared equipment,

x^2="Equipment", if the PSE was caused by system faults and component malfunctions, x^3="External", if the PSE was caused by events originating from outside the system, such as natural events and sabotages, or their possible combinations.

(iii) the type of event, y, which can be y^1="Fire/Explosion", in case of ignition of liquid or gaseous substance, y^2="Gas leak", in case of loss of gaseous material, and y^3="Spill", in case of loss of liquid material;

(iv) the Tier level, z, associated to the event, which can be z^1="1", for PSEs with significant LOPC, z^2="2", for PSEs with moderate LOPC, or z^3="3", for PSEs with minor LOPC.

The taxonomy of the barrier has been built by system experts. The possible types of barriers are w^1 ="Technical", w^2 ="Operational" and w^3 ="Organizational". The second sublevel of the taxonomy assigns the words of the vocabulary to the individual barrier type. Figure 4 shows the developed taxonomy of the barriers, with the number of tokens associated to each barrier type.

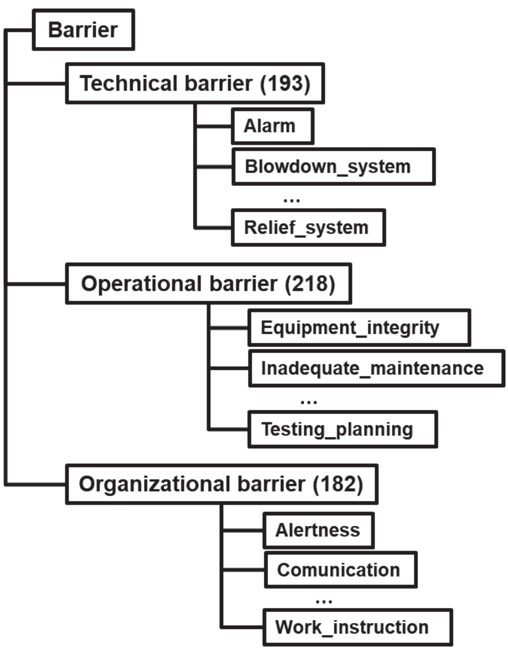

Figure 4: Developed taxonomy for barrier, where the number of tokens assigned to each state is indicated in parenthesis.

Once the taxonomy has been built, the model is completed by defining the causal relations among its elements. With respect to the BN, since the evolution of a single PSE can be influenced by different types of barriers, eight possible states, which account for all the possible combinations of the three barriers "Technical", "Operational" and "Organizational", are defined. Figure 5 shows the obtained BN.

5. Results

The parameters $thresh_{NPMI}$ and $thresh_h$ have been set by trial-and-error to 0.65 and 0.5, respectively. Once the keywords are extracted, the BN model is developed by estimating the parameters of the CPTs according to the procedure described in Section 3.2.2. Finally, the developed BN is used to compute probabilities of interest.
Table 1 reports the unconditional probabilities of the Tier levels, and Table 2 reports the probabilities of occurrence of PSEs of the three Tier levels, conditional to the occurrence of PSEs of the types y^1=Fire/Explosion, y^2=Gas leak and y^3=Spill.

Table 1: Unconditional probabilities of the Tier levels.

Tier level z^c	Probability $P(z^c)$
z^1	0.099
z^2	0.176
z^3	0.724

Table 2: Probability of occurrence of PSEs of the Tier levels z^1, z^2 and z^3, conditional to the type of event.

Probability $P(z^c\|y^b)$	Tier level		
Type of event	z^1=1	z^2=2	z^3=3
y^1=Fire/Explosion	0.048	0.095	0.857
y^2=Gas leak	0.054	0.036	0.910
y^3=Spill	0.120	0.238	0.642

Notice that for the probabilities of occurrence of PSEs with the most severe consequences (Tier level= z^1) in case of event of type y^3=Spill is more than the double of the same probability in case of events of types y^1=Fire/Explosion and y^2=Gas leak. This has been investigated by considering the probabilities of the possible causes, conditional to the occurrence of an event of type "Spill", $P(x^a|y^3)$, $a = 1, \dots, A$ (Table 3).

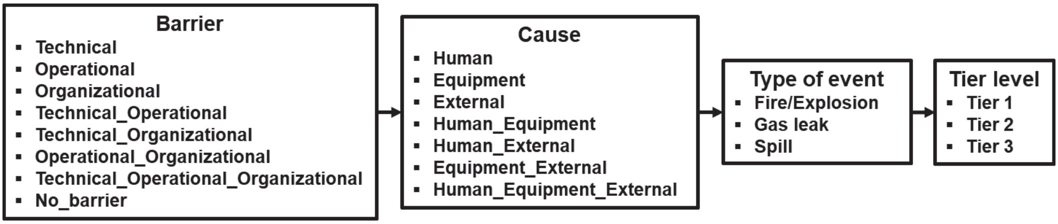

Figure 5: BN used in the considered case study.

Table 3: Probability of occurrence of the PSEs of causes $x^1, \ldots, x^7$, conditional to the type of event y^3.

x^1=Human	0.174
x^2=Equipment	0.695
x^3=External	0
x^4=Human_Equipment	0.131
x^5=Human_External	0
x^6=Equipment_External	0
x^7=Human_Equipment_External	0

The largest contribution involves the cause "Equipment", which describes the occurrence of an equipment failure. To find which type of barrier has mostly influenced equipment failures, we have considered the probabilities of the types of barriers conditional to the occurrence of an event with associated cause "Equipment", $P(w^m|x^4)$, $m = 1, \ldots, M$ (Table 4).

Table 4: Probability of the states of the node barrier, w^m, conditional to the cause x^4 = "Equipment".

Barrier state w^m	Probability $P(w^m\|x^4)$
w^1=Technical	0.078
w^2=Operational	0.205
w^3=Organizational	0.043
w^4=Technical&Operational	0.051
w^5=Technical&Organizational	0.007
w^6=Operational&Organizational	0.081
w^7=Technical&Operational&Organizational	0.023
w^8=No barrier	0.510

It can be noticed that the largest contribution is from the operational barrier (w^2). Considering the reports of type "Spill" and Tier level 1, the most frequent keywords identified by the method and related to the barrier type "Operational" are "lack_of_procedure" and "maintenance". System experts have confirmed that in PSEs of type "Spill" and of Tier level 1, there are issues related to the operational barrier and to procedures implementation when planning and performing inspections and maintenance interventions.

6. Conclusions

A methodology combining a NLP technique and a BN has been developed for the analysis of PSEs in hydrocarbon assets, in support to RA. The methodology allows extracting the knowledge contained in textual reports on the barriers involved during the PSEs and to quantify the probabilities of the severity of the consequences considering different types of events, causes and states of the barriers. The developed methodology has been applied to an electronic repository of reports of PSEs occurred in hydrocarbon plants. The obtained results show that the proposed methodology is able to identify relevant keywords related to barriers that are critical with respect to the occurrence of the PSEs and the severity of their consequences.

References

Ahmed, Saeema, Sanghee Kim, and Ken M Wallace. 2005. "A Methodology for Creating Ontologies for Engineering Design." In *Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference - DETC2005*, Volume 3 B, 739-750. https://doi.org/10.1115/detc2005-84729.

Ansaldi, Silvia M, Paolo Bragatto, Patrizia Agnello, and Maria Francesca Milazzo. 2020. "An Ontology for the Management of Equipment Ageing." In *The 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference*, 978–81. https://doi.org/10.3850/978-981-14-8593-0.

"API RP 754." 2021. American Petroleoum Association (API) Recommended Practice. 2021. api.org.

Bouma, Gerlof. 2009. "Normalized (Pointwise) Mutual Information in Collocation Extraction." In *Proceedings of GSCL 30*, 31–40.

Dayrat, Benoît. 2005. "Towards Integrative Taxonomy."

Biological Journal of the Linnean Society 85 (3): 407–15. https://doi.org/10.1111/j.1095-8312.2005.00503.x.

Decarli, Luca, Luciano Scataglini, Paolo Cherubin, and Laura La Rosa. 2020. "Barrier Integrity Management : An Innovative Approach for Managing Process Safety Risks in Oil and Gas Assets." In *SPE International Conference and Exhibition on Health, Safety, Environment, and Sustainability*. https://doi.org/https://doi.org/10.2118/199472-MS.

Feng, Wenxing, Xiaoqiang Xiang, Guangming Jia, Lianshuang Dai, Yulei Gu, Xiaozheng Yang, Qingshang Feng, and Lijian Zhou. 2012. "Applying the Quantitative Risk Assessment (QRA) to Improve Safety Management of Oil and Gas Pipeline Stations in China." In *Proceedings of the 2012 9th International Pipeline Conference (IPC2012)*, 505–12. https://doi.org/https://doi.org/10.1115/IPC2012-90130.

Hodkiewicz, Melinda, Johan W Klüwer, Caitlin Woods, Thomas Smoker, and Emily Low. 2021. "An Ontology for Reasoning over Engineering Textual Data Stored in FMEA Spreadsheet Tables." *Computers in Industry* 131. https://doi.org/10.1016/j.compind.2021.103496.

Itoh, K, N Omata, and H B Andersen. 2009. "A Human Error Taxonomy for Analysing Healthcare Incident Reports : Assessing Reporting Culture and Its Effects on Safety Performance." *Journal of Risk Research* 12:3-4: 485–511. https://doi.org/10.1080/13669870903047513.

Jensen, Finn V, and Thomas Dyhre Nielsen. 2007. *Bayesian Network and Decision Graphs*. Vol.2. Springer.

Khakzad, Nima, Faisal Khan, and Paul Amyotte. 2013. "Dynamic Safety Analysis of Process Systems by Mapping Bow-Tie into Bayesian Network." *Process Safety and Environmental Protection* 91: 46–53. https://doi.org/10.1016/j.psep.2012.01.005.

Li, Xinhong, Guoming Chen, and Hongwei Zhu. 2016. "Quantitative Risk Analysis on Leakage Failure of Submarine Oil and Gas Pipelines Using Bayesian Network." *Process Safety and Environmental Protection* 103: 163–73. https://doi.org/10.1016/j.psep.2016.06.006.

Mackie, S I, M B Welsh, and M D Lee. 2006. "An Oil and Gas Decision-Making Taxonomy." In *SPE Asia Pacific Oil & Gas Conference and Exhibition*. https://doi.org/https://doi.org/10.2118/100699-MS.

Maio, F Di, O Scapinello, E Zio, C Ciarapica, S Cincotta, A Crivellari, L Decarli, and L Larosa. 2021. "Accounting for Safety Barriers Degradation in the Risk Assessment of Oil and Gas Systems by Multistate Bayesian Networks." *Reliability Engineering and System Safety* 216 (March). https://doi.org/10.1016/j.ress.2021.107943.

Milana, Diletta, Maria Stella Darena, Nico Bettio, Chiara Cerruti, Guido Siliprandi, Alessandra Fidanzi, Paolo Cerioli, et al. 2019. "Natural Language Understanding for Safety and Risk Management in Oil and Gas Plants." In *International Petroleum Exhibition & Conference*.

Parks, Donovan H, Maria Chuvochina, David W Waite, Christian Rinke, Adam Skarshewski, Pierre-alain Chaumeil, and Philip Hugenholtz. 2018. "A Standardized Bacterial Taxonomy Based on Genome Phylogeny Substantially Revises the Tree of Life." *Nature Biotechnology* 36: 996–1004. https://doi.org/10.1038/nbt.4229.

Rohmer, Jeremy. 2020. "Uncertainties in Conditional Probability Tables of Discrete Bayesian Belief Networks : A Comprehensive Review." *Engineering Applications of Artificial Intelligence* 88. https://doi.org/10.1016/j.engappai.2019.103384.

Saidani, Michael, Bernard Yannou, Yann Leroy, François Cluzel, and Alissa Kendall. 2019. "A Taxonomy of Circular Economy Indicators." *Journal of Cleaner Production* 207: 542–59. https://doi.org/10.1016/j.jclepro.2018.10.014.

Sattari, Fereshteh, Renato Macciotta, Daniel Kurian, and Lianne Lefsrud. 2021. "Application of Bayesian Network and Artificial Intelligence to Reduce Accident / Incident Rates in Oil & Gas Companies." *Safety Science* 133. https://doi.org/10.1016/j.ssci.2020.104981.

Valcamonico, Dario, Piero Baraldi, and Enrico Zio. 2021. "Natural Language Processing and Bayesian Networks for the Analysis of Process Safety Events." In *5th International Conference on System Reliability and Safety (ICSRS)*, 216–21. https://doi.org/10.1109/ICSRS53853.2021.9660733.

Vinnem, J E, R Bye, B A Gran, T Kongsvik, O M Nyheim, E H Okstad, J Seljelid, and J Vatn. 2012. "Risk Modelling of Maintenance Work on Major Process Equipment on Offshore Petroleum Installations." *Journal of Loss Prevention in the Process Industries* 25 (2): 274–92. https://doi.org/https://doi.org/10.1016/j.jlp.2011.11.001.

Xiaoguang, G A O, Yang Yu, and G U O Zhigao. 2019. "Learning Bayesian Networks by Constrained Bayesian Estimation." *Journal of Systems Engineering and Electronics* 30 (3): 511–24. https://doi.org/10.21629/JSEE.2019.03.09.

Zeng, Zhiguo, and Enrico Zio. 2018. "Dynamic Risk Assessment Based on Statistical Failure Data and Condition-Monitoring." *IEEE Transactions on Reliability* 67 (2): 609–22. https://doi.org/10.1109/TR.2017.2778804.

Zio, Enrico. 2018. "The Future of Risk Assessment." *Reliability Engineering and System Safety* 177: 176–90. https://doi.org/10.1016/j.ress.2018.04.020.