Article

Preliminary Design and Construction Database for Laboratory Accident

Xuying Zheng[1,2], Fang Miao[3], Jiaqi Yuan[4], Huasong Xia[4], Piyachat Udomwong[1,] * Nopasit Chapitcak[1,*]

1. International College of Digital Innovation, Chiangmai University, Chiangmai 50200, Thailand
2. School of Architecture and Civil Engineering, Chengdu University, Chengdu 610106, China
3. Big Data Research Institute, Chengdu University, Chengdu 610106, China
4. College of computer science, Chengdu University, Chengdu 610106, China
* Correspondence: piyachat.u@cmu.ac.th (P.U.); nopasit@cmuic.net (N.C.)

Abstract: With the growing of university chemistry experiment projects, scientific research personnel, specialized equipment, laboratory accident is increasing yearly. And accident data lacks a safety platform to store related information and cannot be guaranteed for efficient conditional sharing. To solve these problems, we designed a laboratory accident system to store, share related data and predict risk level. In this paper, we manually collected chemistry laboratory accidents by python software and class assignments, then analyses risk factor variables using Spsspro, finally established a prediction model using Stata. We intended to register laboratory related data into proposed chemistry accident system based on data ownership safety architecture. The chemistry accident system can break data barriers using confirmation and authorization key technology to trace non-tampered data source in real time when emergency accident happens. Meanwhile, our proposed system can use our designed accident risk model to predict risk level of any experiment project. It can also recommend appropriate safety education models.

Keywords: chemistry laboratory accident, risk factor analysis, laboratory accident system, manual accident database, quantitative prediction risk

1. Introduction

With the increasing of laboratories, experiments projects, various equipment, number of research persons, diverse method to learn safety class, university chemistry accident is also increased. Chemistry laboratory accident is involved in explosions, chemicals leakage, equipment improper operation which cause high cost and casualties. We usually find accidents from website or literature. The existing accident systems are mainly focused on instruments management, education, safety studies video. Different system is independent, such as chemicals systems, special equipment systems, safety test systems. If we have emergency accident, we need retrieve data, it costs time to trace data source. So, we need to have a system to register safely with techniques, also it can share information in real time when it needed.

Researchers have collected nearly 10 years chemistry university accident from websites [1]. They also collected by literature reviews to analysis which parts more easily cause accidents [2][3]. Some researchers want to analysis one or several typical accidents which wants to get safety guide recommend or formulate safety regulators [4][5][6]. Some emergency departments want to recreate animation of the accident.

Safety education is more and more attention, stem education, block chain [7][8]. However, it seems no expected effect. We need use more quantitative method to recommend suitable safety courses for experimenters.

Existing accident literature has regional accident research key accident case analysis or annual accident analysis [9][10][11]. We lack a whole website laboratories accident collection.

To solve laboratory safety experiment, it has engineering safer systems [12]. A remote lab for the "Data Acquisition Systems", delivered as the digital twin lab [13]. Internet of Things (IoT) used in many laboratories to decrease safety risks connected with equipment in an automated way [14]. It also has engineering program and its associated Smart Lab to develop students capabilities in the areas of Artificial Intelligence (AI) [15].Some researchers analyze the situation of university laboratory management and realizes the necessity of computer technology to manage laboratory. They used the Internet of things to build the university laboratory management system and given credit security evaluation model[16].It has a method that connects an equation model (SEM) with system dynamics (SD) is presented to dynamically assess lab safety with the insufficient data [17].Researchers proposed a real-time smart vision-based lab-safety monitoring system to verify the safety protection of students from video [18].

It needs a system which can store related accident data safely and search data source quickly, meanwhile, it also needs predict risk quantitative. So, our paper proposed a chemistry accident system based on data ownership safety architecture to solve these research gap. Firstly, we collected chemistry accident using python and manual, then used these historical accident data to choose risk variables by Spsspro, and design accident model using Stata. Secondly, we design our data ownership safety architecture accident system which let data owner register safely and conditional sharing quickly with key algorithm. Finally, we put this accident manual database to our proposed systems, then used risk model in systems to predict experiment risk level, it also can recommend suitable education safety class.

Data ownership safety architecture is described one body and two wings which one body is the data combined with ownership. The one wing is that data can register innately, another wing is the data should use the key technology to protect data during conditional sharing [19]. Data ownership safety architecture application has advantage on tourism, smart cities and laboratory management[20].

2. Materials and Methods

2.1 Manual laboratory data collection database

We used lab accident, chemicals accident, equipment accident these key words to search public academic database with python software to get 854 chemistry laboratory accident literature.

Then we took environmental engineering undergraduate safety class assignments on first semester in 2022 which requires 240 students each collected 10 chemistry laboratory accidents from non-academic websites. The manual assignment databases need to collect accident meta information including happened year, season, time, university, major, equipment, chemicals usage, experimenters' degree, whether the teacher is on site or not, lost money, injured and dead number, data source website link which is shown in Table 1.

Table 1. Meta information of laboratory accident.

variable	type	descriptions
years	numerical	laboratory accident happened year
season	character	laboratory accident happened season
location	character	laboratory accident happened university location
experimental classification	character	experimental divides teaching (instructed by teachers); scientific research (research team), factory and university cooperation project
dead	numerical	number of person dead in accident
injure/dead number	numerical	number of persons injured or dead in accident
accident style	character	accident type divided in explosions, chemicals, mechanical injury
degree	character	person in accident who are undergraduate, master or PhD
equipment	character	equipment has common, special, gas, heating type
chemicals	character	chemicals used in experiment

teacher in accident	character	are teaches in accident (yes/no)
accident cause public	character	is the accident reason public (yes/no)
university level	character	university level divided into normal 211 985 by Chinese criteria and university rank
money lost	numerical	money lost in accident
data source	character	which website found these accidents

2.2. DOSA

We intend to put manual collected accident databases to a prediction system. Many research teams do not want to share their ordinary accident to public, so we have difficult to find these minor accidents on websites or literature. They also care about their lab related data privacy, and search data quickly if any emergency happens. So, we choose data ownership safety architecture to design our system.

DOSA is proposed by Miao on 2008, it has data register center (DRC), data authorization center (DAC)and key technology to protect data ownership and privacy, we used it to break data barriers, mining data more value. DAC, we choose suitable key algorithm to guarantee data owner and user unconditional sharing data. We choose SM2 for grade 5 data which need high protection [21].

2.3. Design laboratory accident system method

We used data ownership safety architecture to design a manual database accident system. Our proposed system can register accident-related data safely and allow conditional sharing with key technology. We also can use new experiment projects information to predict risk level.

Data owner can encrypt their accident data into DRC with their own public key to authorization, then system can automatically generate accident category. Though Spsspro analysis results and Stata risk model, a new project can input related data to predict this accident risk level. If the risk level is high, the researcher can get a suitable safety recommendation on proposed system.

3. Results

This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn.

3.1. Laboratory accident data risk analysis

By cleaning collected accident data, we classified equipment into common equipment, special equipment, gas cylinder and non-equipment. Chemicals are divided into common chemicals, hazardous chemicals, and non-chemicals by MSDS classification. Experimenter degree included doctoral, master, undergraduate, college. University level is classified into normal, high, and top-level using university rank documents and Chinese 211,985 documents. After cleaning data, we got 220 chemistry university laboratory accident databases.

We use python to collect laboratory accident-related literature review. The word cloud generated by using python is shown in Figure 1.

Figure 1. Chemistry accident word.

We need to choose key risk variables from collected factors, and decided which factors can be dependent variable. Through historical chemistry laboratory accident databases, we used Spsspro to analysis risk factor.

(a)

(b)

(c)

(d)

(e)

(f)

Figure 2. Analysis chemistry lab accident risk factor including season, experiment type, accident type, equipment, and chemicals.

From Figure 2(a), we can observe accident happens season, spring, summer, and winter is no noticeable difference. The reason that accidents happened on autumn is repetitively minimal is September is new semester opening, university have been do concentrated safety education. According to experiment type, it shows the proportion of teaching experiment is higher than research experiment (Figure 2b). However, the number of attends on scientific is less than the joined teaching experiment students. From percentage of experimenters, the researcher experiments are more easily happened accidents (Figure 2c). According to chemicals to accidents effect, hazardous chemicals is most serious. Common chemicals ratio equals to gas usage; however, chemistry experiments are less use gas. Gas is more danger than common chemicals. Electric harm, injection, and mechanical injury account a smaller proportion (Figure 2d). From collecting all accident types, the proportion of the explosion is as high as 85%, and the chemicals leakage ranked number 2. Other reasons are all around 1%. The reason is particularity of chemistry laboratory, it used hazardous chemicals and special equipment do experiments can cause lager explosion which causes injured or economic losses. Chemicals leakage can cause long-term environment pollution which also need to be attention (Figure 2e). It shows common equipment account 75% in accident, and special equipment about 20% which includes high pressure equipment, gas, etc. And the infrastructure involves water, electric, exhaust architecture is nearly 3% in collected database (Figure 2f). We can see teachers or staff not in laboratory is 2 times easier to happen an accident. In figure 3, it is shown recently is more and more accident happened.
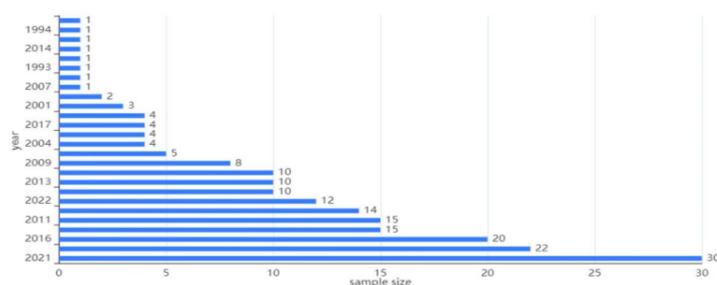

Figure 3. Chemistry laboratory accident numbers in happened year.

Then we used above analysis results, we have eliminated season variables. And we co minded injured, dead number with accident type into accident level which from class1 to class 5. We choose six factors to make a heat map using Spsspro which is shown in Figure 4.

Figure 4. Chemistry laboratory accident heat map.

From Figure 4 we can observe, Q6 experiment type is relatively strong correlation with the Q10 experimental accident level. Q8 accident injured number is negative correlation with Q10 experimental accident level. And accident level is positive correlation with use of Q18 chemicals usage. Whether the teachers are in lab or not is also relative with accident level and experiment type. So, next step we used these analysis results to choose variables and establish accident model.

3.2. Accident risk model using Stata

Through the above analysis, we choose experimenters degree, experiments used equipment and chemicals as independent variable. And dead, injured number combined with accident type set the causative variable. We used these factors to establish risk model equation using Stata. The regression results by Stata are shown in Figure 5.

```
regress ACCIDENTlevel specialequipment gas hazardouschemicals
```

Source	SS	df	MS			
				Number of obs	=	196
				F(3, 192)	=	8.15
Model	40.8491147	3	13.6163716	Prob > F	=	0.0000
Residual	320.844763	192	1.67106647	R-squared	=	0.1129
				Adj R-squared	=	0.0991
Total	361.693878	195	1.8548404	Root MSE	=	1.2927

ACCIDENTlevel	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
specialequipment	.7265465	.2737966	2.65	0.009	.1865111	1.266582
gas	.8661627	.3738602	2.32	0.022	.1287621	1.603563
azardouschemicals	.499994	.2382289	2.10	0.037	.0301121	.9698759
_cons	1.976153	.2116744	9.34	0.000	1.558647	2.393659

Figure 5. Risk model established using Stata.

From regression model, we choose college student, non-equipment, and non-chemicals as benchmark. Because part of data is lacked, Stata certified sample size is 196.From Stata results, we can see person degree, common chemicals and infrastructure have no strong correlation for accidents. However, hazardous chemicals, gas and special equipment have relationship with chemistry laboratory accidents took them as variables, so the model equation is:

$$AL=a+b*SE+c*G+d*HC \tag{1}$$

t   (2.65)(2.3)(2.1)

AL=accident level SE= Special Equipment G=Gas HC=Hazardous Chemicals

a=1.976153 b=0.7265465 c=0.8661627 d=0.499994

3.3. Design laboratory accident data system

Though manual collected database quantitative analysis, we choose strong and weak connection variables，then design whole chemistry laboratory accident structure.

We designed to have a system to store all related laboratory accident data into DRC, including consumables broken, glass consumables heating explosion, experiments explosion, etc. If an accident happened, we need to store related data, such as accident time, location, lab, experimenters, projects, chemicals, equipment, consumables, real time IOT photo and video into DRC with SM2 or AES key algorithm to authorize data ownership.

Then we can link important laboratory accident-related data according to above analysis using MySQL which is shown in Figure 6. If an accident happened in a laboratory,

we could search from Lab-id to trace person who charge of lab and project supervisor information. We also can search this project used equipment; application chemicals record to design emergency rescue plan. Though equipment-id, we can find equipment-details, including bidding contract, companies, use manual, etc. We can combine IOT video to judge whether the operator or equipment caused the accident. Using lab-id we can search laboratory infrastructure, such as water pipe material, circuit maintenance records. Meanwhile, IOT equipment can real-time monitoring which also need encrypted upload to proposed system.

These data are encrypted to register laboratory accident system based on DOSA. Only emergency happens, data owner use key algorithm to authorized emergency department which can break isolated data islands.
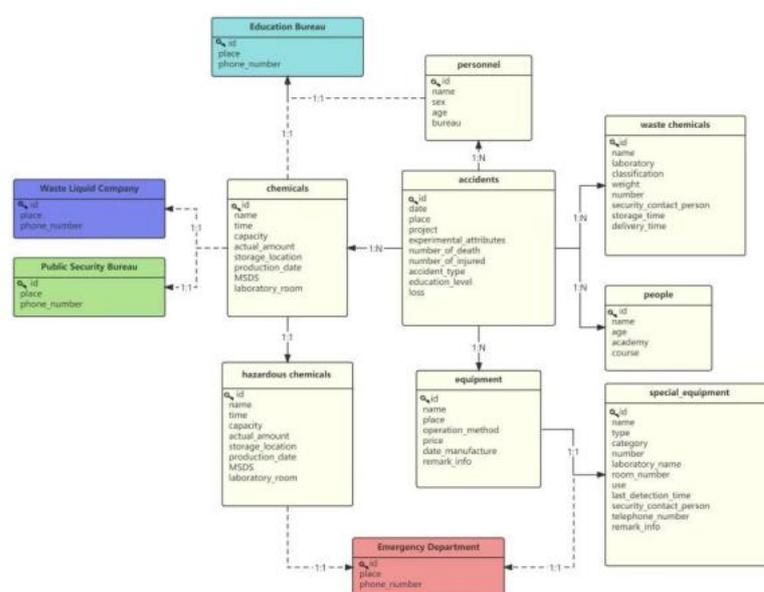

Figure 6. Structure of laboratory accident data.

Supervision departments needs to search accident-related data, such as hazardous chemicals collection. Each laboratory data owner can find supervision departments public key from DAC, then encrypted chemicals data to departments' public key. Supervision department can use their private key to decrypt data. Table 2 presents the function and meaning. The process of encrypted accident data to DRC using SM2 is shown in Algorithm 1.

Table 2. Function and meaning.

Function	Meaning
connect_to_mysql	Connect to MySQL database
open_excel_file	Open Excel file
select_sheet	Select a worksheet in an Excel file
read_data	Read data in Excel file
sm2_encrypt	Encrypt cells using SM2 algorithm
construct_insert_sql	Build SQL insert statement
execute_sql	Execute SQL statement
close_mysql_connection	Close the MySQL database connection

Algorithm 1. Sm2enc_to_mysql ()

    Input: excel of accident data

    Output: None

```
1.    db = connect_to_mysql(database, username, password)

2.

3.    excel_file = open_excel_file("accident_data.xlsx")

4.    sheet = select_sheet(excel_file, "accident")

5.    accident_data = read_data(sheet)

6.

7.    encrypted_data = []

8.    For row in data:

9.        encrypted_row = []

10.       For cell in row:

11.           encrypted_cell = sm2_encrypt(cell)

12.           encrypted_row.append(encrypted_cell)

13.           encrypted_data.append(encrypted_row)

14.       End for

15.   End for

16.

17.   accidents = " accident_table"

18.   columns = ["column1", "column2", "column3"]

19.   For row in encrypted_data:

20.       values = []

21.       For cell in row:

22.           values.append(cell)

23.           sql = construct_insert_sql(accidents, columns, values)

24.           execute_sql(db, sql)

25.       End for

26.   End for

27.

28.   close_mysql_connection(db)
```

4. Usage of laboratory accident system

This web designed method is data ownership safety architecture. Data owner can register this system which provide a key pair. Our system can transmission private key to data register and reserve the public key to DAC. If the lab data owner wants to register accident data to this system, he can use his public key to encrypt this data to confirm ownership, meanwhile data category can automatic generated. Data user can search category to find target data which need apply to the data owner or DIY data product.

Figure 7shows our proposed system can link related data sources and data owners, such as regulator departments, chemicals and waste liquid companies, various lab, different research teams. These data owners can encrypt their data to this system with their own public key which can confirm data ownership.
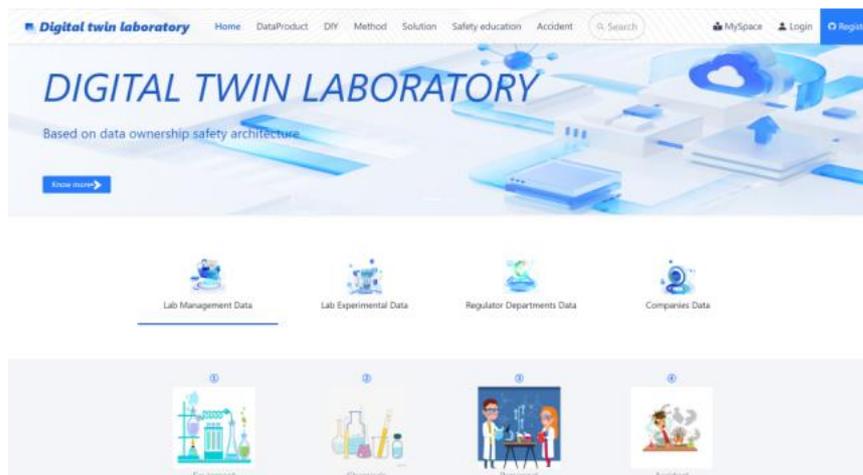
Figure 7. Main lab accident-related data web.

Figure 8 shows our system can collect historical lab accident or happening accident. The research team do not want to public their laboratory accident, they can encrypt their accident into system with their own public key which even administrator cannot view this information without permission. But if they have emergency or supervision, they can use supervision public key to encrypt which can authorization.

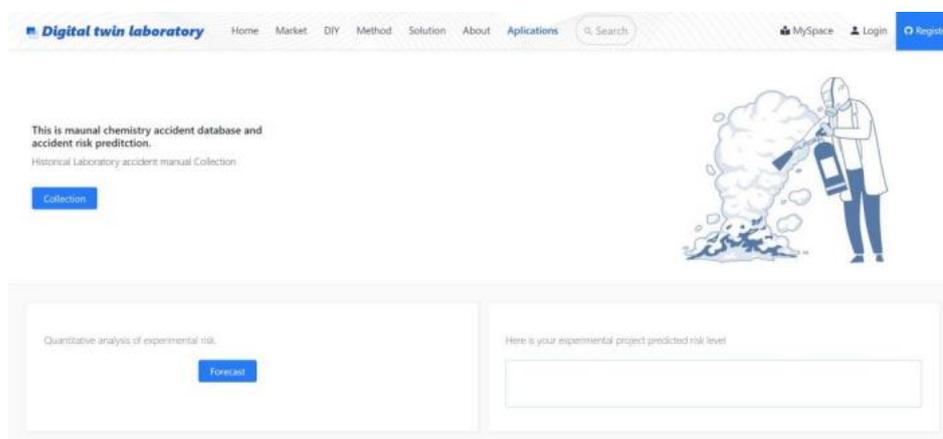**Figure 8.** Historical laboratory accident data collection.

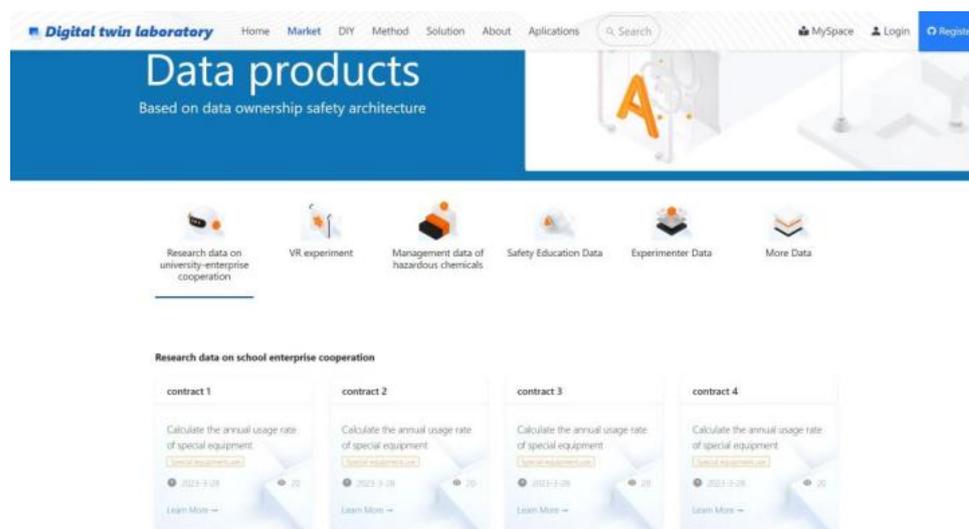**Figure 9.** Prediction of lab accident risk.

Figure 10. Data product such education safety class.

Figure 9, if any new project will do experiments in lab, they can put their data using our risk equation to predict level. Figure 10, if the experimenter's predicted risk is relatively high, it can choose VR experiment or recording to learn safety class.

5. Discussion

This system firstly can link all laboratory accident-related data which can break data barriers. Secondly, it provides quantitative prediction model which will decrease accident probability and improve safety learning efficiency. Thirdly, it based on data ownership safety architecture using key algorithm. So, if any emergency accident happens, we can quickly trace related data using this system. It can provide real-time and non-tampered data to rescue team. Finally, we also can extend this system function and apply for other occasion, such as smart cities.

6. Conclusions

In this study, we developed a chemistry laboratory accident system, we have used prediction model which analysis by collecting historical accident data. The proposed system can guarantee data owner's interest, meanwhile it breaks accident-related data barriers. Our contribution is manual organize accident data using Spsspro to choose significant variable, then establish an accident level predication model. Second contribution is we designed an accident system based on data ownership safety architecture which data owner and user can conditionally share data with key algorithm. This method guarantees safety and efficiency if an emergency accident happens. Third contribution is we combined risk model into our proposed system, any experiment project wants to attend in laboratory, it can put related data to predict this project risk level. If the level is high, the research team can study suitable safety education in proposed system which can decrease the accident risk.

References

1. Eighmy, T., Schovanec, L., Young, A., Martin, J., & Casadonte Jr, D. J. (2020). Ten Years After the Texas Tech Accident: Part I: A Historical Retrospective. ACS Chemical Health & Safety, 27(2), 105-113.
2. Fukuoka, K., & Furusho, M. (2022). A new approach for explosion accident prevention in chemical research laboratories at universities. Scientific reports, 12(1), 3185.
3. .Ménard, A. D., & Trant, J. F. (2020). A review and critique of academic lab safety research. Nature chemistry, 12(1), 17-25.
4. Ali, L., Alnajjar, F., Parambil, M. M. A., Younes, M. I., Abdelhalim, Z. I., & Aljassmi, H. (2022). Development of YOLOv5-Based Real-Time Smart Monitoring System for Increasing Lab Safety Awareness in Educational Institutions. Sensors, 22(22), 8820.

5. Nasrallah, I. M., El Kak, A. K., Ismaiil, L. A., Nasr, R. R., & Bawab, W. T. (2022). Prevalence of accident occurrence among scientific laboratory workers of the public university in Lebanon and the impact of safety measures. Safety and health at work, 13(2), 155-162.

6. Allen, K. (2014). A young lab worker, a professor and a deadly accident. Toronto Star.

7. Walters, A. U., Lawrence, W., & Jalsa, N. K. (2017). Chemical laboratory safety awareness, attitudes and practices of tertiary students. Safety science, 96, 161-171.

8. Love, T. S., Roy, K. R., & Sirinides, P. (2023). A national study examining safety factors and training associated with STEM education and CTE laboratory accidents in the United States. Safety Science, 160, 106058.

9. Miller JD (2006) Ohio cites university in death. The Scientist Magazine®

10. Mohamad Z (2012) Laboratory quality management requirements of engineering at the polytechnics Ministry of Higher Education Malaysia

11. Foderaro LW (2011) Yale student killed as hair gets caught in lathe. The New York Times

12. Leveson, N. (2004). A new accident model for engineering safer systems. Safety science, 42(4), 237-270.

13. Georgakopoulos, I., Piromalis, D., Ntanos, S., Zakopoulos, V., & Makrygiannis, P. (2023). A Prediction Model for Remote Lab Courses Designed upon the Principles of Education for Sustainable Development. Sustainability, 15(6), 5473.

14. Betharajoo, S. K., Paramasivam, S., & Chua, H. S. (2021). Safety Risk Assessment and Warning System at Engineering Laboratory Through IoT Technologies. In Advances in Electrical and Electronic Engineering and Computer Science (pp. 13-23). Singapore: Springer Singapore.

15. Zamiri, M., Sarraipa, J., Ferreira, J., Lopes, C., Soffer, T., & Jardim-Goncalves, R. (2023). A Methodology for Training Toolkits Implementation in Smart Labs. Sensors, 23(5), 2626.

16. Wei, C. (2020). Research on university laboratory management and maintenance framework based on computer aided technology. Microprocessors and Microsystems, 103617.

17. Ma, L., Ma, X., Zhang, J., Yang, Q., & Wei, K. (2021). A Methodology for Dynamic Assessment of Laboratory Safety by SEM-SD. International Journal of Environmental Research and Public Health, 18(12), 6545.

18. Ali, L., Alnajjar, F., Parambil, M. M. A., Younes, M. I., Abdelhalim, Z. I., & Aljassmi, H. (2022). Development of YOLOv5-Based Real-Time Smart Monitoring System for Increasing Lab Safety Awareness in Educational Institutions. Sensors, 22(22), 8820.

19. Miao, F., Yang, W., Ye, A., & Chen, H. (2016). The architecture for data security and application in Digital Earth platform. In IOP Conference Series: Earth and Environmental Science (Vol.46, no. 1, p. 012050).

20. Zheng, X., Miao, F., Chakpitak, N., Huang, J., & Wang, J. (2021, December). Discussion of University Chemistry Laboratory Management Using DOSA Platform and Safety Education Based on Blockchain. In 2021 4th International Conference on Education Technology Management (pp. 79-85).

21. Zheng, X., Miao, F., Udomwong, P., & Chakpitak, N. (2023). Registered Data-Centered Lab Management System Based on Data Ownership Safety Architecture. Electronics, 12(8), 1817.