

Requirements for Central Repository: Flood Management Malaysia

Mohd Ismawira Mohd Ismail, School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia, wira@i-erat.com.

Mohammad Nazir Ahmad, Institute of Visual Informatics, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia, mnazir@ukm.edu.my.

Nor Hidayati Zakaria, Azman Hashim International Business School, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia, hidayati@utm.my.

Razatulshima Ghazali, ICT Consulting, Malaysian Administrative Modernisation and Management Planning Unit, Cyberjaya, Malaysia, razatulshima@mampu.gov.my.

Article Info

Volume 82

Page Number: 11908 - 11915

Publication Issue:

January-February 2020

Abstract

Abstract: Flood Management (FM) is a combined effort used to tackle flooding. In Malaysia, it involves multiple government agencies, NGOs and individuals who respond to flood related incidents. Knowledge sharing activities need to be governed and disaster-related knowledge shared among the major players must be systematically managed, especially in this digitized society. The crucial mechanisms to enable knowledge sharing among these players are still loosely defined. From a knowledge management process perspective, mainly on the knowledge storage process, we find it important to propose a Central Repository (CR) for supporting disaster management interoperability. Designing this CR is important. Unfortunately, no clear requirements for its design have been reported in any available literature to date. Therefore, this paper attempts to establish key requirements for a CR for FM. The CR can also be used to facilitate and store lessons learned from previous flood events, to be used in any future events as reference. Currently, FM in Malaysia has limited availability of information as data is scarce.

Article History

Article Received: 18 May 2019

Revised: 14 July 2019

Accepted: 22 December 2019

Publication: 21 February 2020

Keywords: Big Data, Central Repository, Central Repository Design, Flood Management, Interlocking Institutional Worlds, Knowledge Sharing, Open Data.

I. INTRODUCTION

Malaysia has experienced some natural disasters over the years and among these, floods have been identified as a major risk. Many positive studies and efforts has been conducted separately, resulting in relevant knowledge and data for reference, which can be seen as a real combined effort against flooding [20], [28], [29]. Disasters usually bring together a diverse set of organizations who all play supportive roles to assist in the disasters, and this in turn produces a large amount of data, which must be managed to potentially improve efficiency for future incidents [23], [24]. ICT can effectively support this albeit there has been lack of effective ICT solutions to

Support the coordination efforts. This is also the case in Malaysia where relevant research, Standard Operating Procedures (SOP), policies and data on disasters are kept at the sectoral level resulting in slow learning curves and repetitive mistakes in handling the situation [16].

The studies mentioned above highlighted a few key points, such as indicating that floods are a domain involving many stakeholders. These stakeholders require systematic data collection and analyses to make informed policy decisions and they also need to have an effective information distribution platform which does not hamper information flow to the

society. A central platform is required to support this kind of domain, especially when it involves such a wide range of stakeholders. These stakeholders and their interactions usually generate new valuable data. Thus, a strategy must be developed to efficiently organize and manage this data. The proposed platform would govern the processes of participating stakeholders, and thus, technical and quality standards must be established and maintained.

However, the Sendai Framework Data Readiness Review Report 2017 on Malaysia by the United Nation Office for Disaster Risk Reduction (UNISDR) indicated that Malaysia does not have a national database for collecting disaster related knowledge. Additionally, there are no well-defined key requirements for such a national database or in other words, there is no national disaster Central Repository (CR) for Malaysia. Thus, this paper mainly aims to address a central repository de-sign requirement for flood management (FM) in Malaysia.

II. BACKGROUND

a. A Centralized Platform for a Domain like Flood Management in Malaysia

Floods are considered to be Malaysia's most devastating natural disasters. Despite experiencing numerous disastrous floods, there has not been much progress in the areas of lessons learned from these floods, to be used as knowledge for any such events in future [9]. It is believed that relevant research, Standard Operating Procedures (SOP), policies and data on floods are available but are held by sectoral agencies. It is therefore crucial to establish a CR for Flood Management in Malaysia.

b. FM as Interlocking Institutional Worlds Domain

FM involves many stakeholders who share common interest to overcome the flood issues. The need to communicate, interoperate, share and exchange information in this kind of domain has been addressed in the Interlocking Institutional Worlds (IWs) concept as introduced by [6]. The root concept was derived from philosophical studies of the General Theory of

Institutional Facts [25] and the Speech Act Theory [2]. The concept explains key messages such as the phenomenon of interoperating institutions where their information systems interoperate, and the important role of ontology (conceptual model) as a specification of the institutional world (institution's conceptualization) to support institutions' information systems to interoperate. Some of the domains adopting these can be seen in the social reality of semantic web, Olympics, Postal codes, Map preparation and Geographical Information Governance [7], Flood Management [17], Halal Food Management [13] and Les-sons Learned Management [3].

c. ICT in Flood Management

ICT plays a significant role in managing disasters such as floods. ICT is useful and applicable to each of the disaster management phases which include mitigation, preparation, response and recovery. The importance of ICT in flood management has been well documented as ICT facilitates information sharing [19], helps in synergizing flood management efforts [10] and acts as an agent to manage continuous change [1].

d. Central Repository

Generically it is a central location for data lifecycle management. It acts as shared resources for users to capture valuable information, host the information, distribute and reuse that information for greater benefit. The most common type of repositories is centralized and distributed. Centralized repository is a single repository where all users access it to perform their activities while distributed repositories are normally split across different location in the network. There were many proven benefits of having a central repository to enhance productivity, efficiency and decision making.

e. Requirement for a New CR Design to Support IWS Domain like Environments

The obvious problem in the current literature is related to design issues of CR. It can be clearly seen from the existing literature, that most of the CR

designs are influenced by data warehouse development projects. This is because the nature of a CR is to gather data from multiple sources and organize them for new demand or reuse. Traditionally, data warehouse developments were inspired by data-centered views (bot-tom-up approach of integrating databases model) which are similar to the ones seen in federated databases research studies. The research ceased in the mid-1990s because it was discovered that two systems could in general not be federated due to semantic heterogeneity [6]. Therefore, we need to introduce a different way of designing a CR which does rely on data, but with a primary purpose of supporting the stakeholders.

The CR to be developed is to be action-centered in supporting interoperation in-stead of something strictly data oriented. We need to define key general requirements for the CR design. However, there are very limited literature sources that look into well-defined requirements for designing a CR platform which is primarily action-centered. Most literature tends to focus on the technology framework and data-centered view, and ends up neglecting the business processes when designing the plat-form.

III. METHODOLOGY

In this paper, our method to define the general requirements for the CR design is di-vided into three (3) stages. The fourth stage which is the field work will not be ad-dressed in this paper. Fig. 1 below depicts our method in deriving the eight requirements.

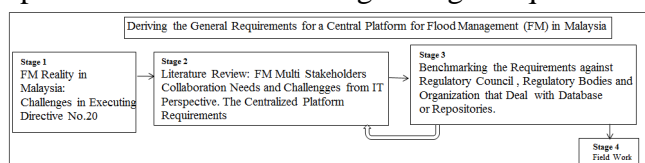


Fig. 1. Methodology Flowchart

a. Stage 1

In the first stage we look into the reality of FM, where the tasks can be classified into four phases which are mitigation, preparedness, response and recovery, and guided by the Directive No.20. All phases relate to the capability of handling emergency tasks and require the stakeholders to interconnect in

activities to coordinate their efforts and resources [4].

FM in Malaysia faces the same challenges as other countries as indicated by the Federal Emergency Management Agency (FEMA) where it highlighted the following:

- The size of response efforts can be extremely large.
- Complexity in interactions can be seen. Many reports with different perspectives are available, but no single snapshot exists that could provide a comprehensive picture.
- There have been reports on increased inter-agency communication in extreme disaster events.

Based on the review, it can be said that there is a need to have a mechanism to improve the overall knowledge sharing capabilities in FM. At this stage, three (3) requirements have been identified as important for the design of a CR for FM in Malaysia, namely, centralization, standardization and service process oriented.

Centralization is the integral key requirement in the design of the CR for FM due to the reason that the FM domain has many stakeholders, different applications and various sources. The development of a common repository and its importance in national risk assessment is also something that has been acknowledged by the United Nations Office for Disaster Risk Reduction (UNISDR) [26].

For centralization to work, standards must be established to ensure that all the CR participants are governed when using the CR. The UNISDR has highlighted plans to specify preferred formats which underline the standards in creation, maintaining and sharing of data, to alleviate the compatibility challenges that have been faced in the past. Other groups which also highlight the importance of standards include the German Institute of Standardization [5].

The third requirement, for the CR to be service process oriented, is deemed necessary as most of repositories developed are usually data oriented. This method puts data as primary, but we plan to develop

the CR based on service process as the primary focus. To summarize, Stage 1 confirms the complexity of FM as it involves many agencies with different backgrounds, experiences, applications and processes. Development of a CR based on the review done, will help to consolidate that information, hence making it more effective to learn from experience and assist in decision making to respond quickly to new scenarios.

b. Stage 2

Here we perform a literature review to further understand how different agencies collaborate in joining efforts in the disaster management area. We look deeply into theories that discuss on repository design, knowledge management and inter agency collaboration. Overall the main theme here is about how to govern the stakeholders' participation in the CR design.

In FM, these activities may involve information from many sources. These data need to be integrated by linking them together to produce a common understandable output. Based on this, from the literature review, we have identified additional requirements for the CR which include discussions on big data, open data, knowledge sharing, public and private data and publishing and subscription activities. These requirements can be seen as tightly integrated between each other and should be discussed in the design of the CR.

Big data is defined as a collection of datasets that are very large and complex to process especially by using classical database management tools. Big data has opened a new way to look at disasters because of the possibilities to visualize, analyse and predict disasters [27].

Open data is an ongoing discussion as many organizations still face problems in accessing data sources. Open data is often required to provide detailed information, especially to retrieve government data, as it provides important parameters in flood damage assessment [8].

Data shared can be both private and public data, where some control is required as a form of trust to the data owners and users. Knowledge sharing is vital

towards government transformation as it shows the statement of intent from the government to improve services towards the public. Thus, through the CR, the stakeholders will be able to perform publishing and subscription activities.

To conclude, the contribution of Stage 2 here is to make available the body of knowledge on context for the CR that can be instantiated and used as a starting point for consideration in the repository design, especially in governing the stakeholders' participation.

c. Stage 3

In this stage we try to benchmark our requirements for CR design found in previous two stages against repository council, authorization or regulatory bodies and organizations that deal with databases or repositories. The idea is to ensure that the requirements identified will help shape the CR design. The requirements and discussions involve: (1) National Science and Technology Council (NSTC), (2) Primary Trustworthy Digital Repository Authorization Body (PTAB), (3) Federal Agencies Digitization Guidelines Initiative (FADGI), (4) Joint Information Systems Committee (JISC), (5) GOOGLE, (6) Digital Repositories Infrastructure Vision for European Research (DRIVER), (7) Consultative Committee for Space Data Systems (CCSDS), (8) ORACLE, and (9) Irving Fisher Committee (IFC).

Centralization as the first requirement has been adapted by GOOGLE and FADGI. GOOGLE chooses a centralized repository as it supports (1) Unified versioning, one source of truth, (2) Extensive code sharing and reuse, (3) Simplified dependency management, (4) Collaboration across teams and (5) Flexible team boundaries and code ownership. While FADGI explains that a central repository is important to pre-serve, manage and protect knowledge and its associated metadata for long term viability.

Standards are also an important CR design requirement, discussing on interoperability between different agencies, systems and datasets that require consensus on standards to foster agreements on

common approaches and meanings. PTAB, JISC, GOOGLE and CCSDS agree on this.

Service Oriented has not been discussed in detail but it is one of the differentials in our proposed CR design.

The CR promotes Collaborative work and Knowledge Sharing as highlighted by NTSC, FADGI, JISC and DRIVER. The CR also supports administrative processes, collaboration, research, teaching and learning.

Supporting the Open Data initiative and being Big Data ready are important CR design considerations as addresses by FADGI, CCSDS, ORACLE and IFC.

Managing Public and Private data is an important discussion. Generally, it is an accepted principle of data management. It includes accountability, protection, integrity, compliance, disclosure and transparency amongst others as adapted by NTSC, FADGI, JISC and DRIVER.

Lastly, one of the business requirements of the CR is that it should support subscription and publishing facilities, where it includes business intent, usage and owner-ship as highlighted by NTSC, FADGI, JISC and DRIVER.

Overall, we were satisfied with the eight requirements as they were also discussed and considered positive by the authority bodies. The next step in our method is to present the requirements to the agencies involved in FM in Malaysia.

d. Stage 4

As this stage involves getting feedbacks and responses from the stakeholders, which is mostly field work, it is out of the scope of this paper. The results of the field work will be discussed as future work.

IV. REQUIREMENTS

The requirements for the CR for FM were derived as discussed in the methodology section and backed by studies shown in Table I below, addressing the need for in-formation sharing and learning platforms in the Malaysian FM domain.

Table I. Current practice of NSC Directive in Malaysia as compared to COBIT, Hyogo and Sendai Frameworks (Adapted from [21])

Key concepts for examined practices	Present Practice		COBIT	Hyogo	Sendai
	NSC	KSOP			
Accountability	√	√	√	√	√
Communication	√	√	√	√	√
Community involvement	X	√	X	√	√
Information sharing	X	√	√	√	√
Performance measurement	X	X	√	√	√
Learning and growth	X	X	√	√	√
Strategic alignment	√	√	√	√	√
Risk management	√	√	√	√	√
Resource management	√	√	√	√	√
Value delivery	√	√	√	√	√
Institutional support	√	√	√	√	√
Stakeholder engagement	√	√	√	√	√
Cross-agency collaboration	√	√	√	√	√
Education, partnership and research	X	√	X	√	√
Organizational structures	√	√	√	√	√

The stakeholders in Malaysian FM domain consist of the government agencies, NGOS, academicians and individuals. These stakeholders need to interact and share information between them during a disaster event [11]. For instance, the National Disaster Management Agency (NADMA) manages flood disaster events. NADMA must engage with other agencies such as METEOROLOGY department, JPS and police to manage the disaster effectively. Based on our understanding, most of these agencies have their own applications. However, these applications do not have the ability to integrate with each other. With this major limitation, FM in Malaysia is unable to provide detailed records of each previous disaster case. This justifies the need of a central platform. To start with, we need to identify the general and basic CR requirements. As per the findings in the methodology section, we briefly discuss the 8 requirements deemed necessary for the CR design.

a. Central Repository

In FM, data can come from many sources from different agencies. Without a CR, data is saved in different independent places and having to integrate them to obtain a full portfolio is very challenging [18]. Consolidating the data into a central repository will provide flood management a significant edge as it allows the flood authorities to obtain up to date and aggregated data.

b. Standard

When discussing on the complexity of flood management, one cannot deny the need to have a CR to flexibly tie all the disparate resources from all the agencies into a trusted source. Furthermore, a CR can govern FM stakeholders through standards. Standards can improve overall FM response time and reduce costly errors in a flood management context. Standardization defines a platform through which work rules, policies, and operating procedures are formalized and followed [14].

c. Service Oriented

Service oriented architecture requires a complete rethink of the way we conceptualize and structure the flood management information systems. Normally it is an on-going process that needs commitment from major participants. Many developed countries including the US and Australia are transforming their government towards citizen centered and result oriented structures and applying technologies based on service-oriented architecture. Being able to learn and improve in flood management through past experiences and being able to execute appropriate actions in response to disasters would be beneficial in solving issues in their respective domains [12].

d. Support Collaboration Work for Knowledge Sharing

Knowledge sharing is vital in many things and it affects every one of us. In the event of a flood disaster, knowledge sharing of how to handle an emergency process effectively may save lives. During disasters, information flows at many levels and across different agencies. In disasters, information flow is a process that invokes actions and decisions in order to manage the disaster impact [7]. Having a CR can facilitate effective information management and provide a platform that shares all information about the disaster, which in turn leads to quicker responses, effective planning and reducing loss [11].

e. Support Open Data Initiative

The idea behind open data is to make data freely

available to everyone. Open data sources are mostly untapped and they have huge potential. Open data can be seen as a new trend in the world of Information and Communication Technology (ICT) [15]. In Malaysia the adoption of Open Data is slow but Malaysia may improve on this by making two changes, which are, by getting new open data advisors and by partnering with the Open Data Institute.

f. Big Data Ready

Government agencies across all sectors accumulate, collect and analyse many sources of data. Data collected can be in the form of formal data and informal data, and this data as termed by Gartner has characteristics of 4 V's in volume, velocity, variety and veracity. There are many crucial activities where big data plays an important role in helping governments across the globe to solve specific issues [22]. The CR for flood management must be able to cater data from different agencies coming from multiple data sources, as all this data offers great potential.

g. Public and Private Data

Although the Malaysian flood management domain involves various players many of these are still reluctant to share their data. We believe that the delineation of data into private data and public data can be implemented to address the issue.

Public data can be defined as data that can be freely uses and distributed by any-one, including agencies involved in the flood management, with no legal conditions attached.

Private data can be defined as data that has sensitive value and which is associated to a particular agency in flood management, or which is an individuals' data that is personal. This data can only be accessed strictly by those with permissions, ensuring fair processing, minimal intrusion and limited usage purposes.

h. Support Publish and Subscribe Data Activity

As data grows and new sources of data exist, the business expectations also grow in terms of how this data can be analysed faster. In FM, multiple parties

must cater to contribute content and collaborate effectively to support publishing and subscription activities, such as following:

Push. The system upon sensing the occurrence of a specified event automatically sends specific information to another information system. In this protocol a client will actively make itself available to the server to accept new events or updates. The server will push the new updates.

Pull. The system upon sensing the occurrence of a specified event, automatically requests specific information from another information system. In the pull protocol, the client periodically connects to the server and disconnects after each procedure. The client will pull new updates from the server.

V. CONCLUSION

In this paper, we have defined key requirements as presented in Section 4, for de-signing a CR for the purpose of supporting the domain of interlocking institutional worlds such as Flood Management. The requirements defined are derived from our understanding of real-world scenarios, literature reviews and applied steps of reviewing the literature as explained in Section 3. Although the key requirements for a CR here are not final, we are confident that these requirements are good enough to start exploring the design issues of CRs for FM. Previous research on CRs has not addressed majority of these requirements, which has led to our first effort here, to explore a new area in designing a CR platform. In the future, we will confirm all the eight requirements with the fieldwork study involving key stakeholders in the context of Malaysia.

VI. REFERENCES

1. Akhtar, N. (2016). ICT-based learning organization support system: managing continuous change. *Economic and Social Development: Book of Proceedings*, 45.
2. Austin, J. L. (1975). *How to do things with words*. Oxford university press.
3. Badr, K. B. A. (2016). *An ontological approach to conceptual modeling of lessons learned management* (Doctoral dissertation, Universiti Teknologi Malaysia).
4. Borders, B. A. O. (2015). *Interorganizational Collaboration in Crisis Response Management*. Lund University, Sweden.
5. Blind, K., Jungmittag, A., & Mangelsdorf, A. (2011). *The economic benefits of standardization*. DIN German Institute for Standardization. http://www.din.de/sixcms_upload/media/2896/DIN_GNN_2011_engl_akt_neu.pdf and *Summary of Results*: http://www.din.de/sixcms_upload/media/2896/economic_benefits_standardization.pdf.
6. Colomb, R. M., & Ahmad, M. N. (2010). A perdurant ontology for interoperating information systems based on interlocking institutional worlds. *Applied Ontology*, 5(1), 47-77.
7. Dahlan, A. R. A., Dahan, H. M., & Saman, M. Y. M. (2014). *The Success Factors in Government Information Sharing (Gis)—Empirical Findings from the Malaysians Natural Disaster Management Context*. *Information and Communication Technology for The Muslim World (ICT4M)*, 2014 The 5th International Conference on: IEEE, 1-7.
8. Eem, S. H., Yang, B. J., & Jeon, H. (2018). Simplified methodology for urban flood damage assessment at building scale using open data. *Journal of Coastal Research*, 85(sp1), 1396-1400.
9. Fan, C., & Mostafavi, A. (2018, April). *Establishing a framework for disaster management system-of-systems*. In *2018 Annual IEEE International Systems Conference (SysCon)* (pp. 1-7). IEEE.
10. Garba, E. J., & Ahmadu, A. S. (2015). *ICT implementation framework for integrated national security, emergency and disaster management for Nigeria*. *West African Journal of Industrial and Academic Research*, 15(1), 27-37.
11. Hashim, K. F., & Ahmad, M. (2013). *A Theoretical Knowledge Transfer Model in Disaster Management: A Case Study of Managing Floods in Malaysia*, ISICO.
12. Hofmann, M., Betke, H., & Sackmann, S. (2015). *Process-Oriented Disaster Response Management: A Structured Literature Review*. *Business Process Management Journal* 21(5), 966-987.
13. Ibrahim, A. A., Colomb, R. M., & Ahmed, A. H. (2015). *Insuring halal food interlocking institutional worlds integrity using an ontology server*. *International Journal of Computer Science and Telecommunications*, 6(3), 11-17.
14. Jang, Y., & Lee, J. (1998). *Factors Influencing the Success of Management Consulting Projects*. *International Journal of Project Management* 16(2), 67-72.

15. Kassen, M. (2013). A promising phenomenon of open data: A case study of the Chicago open data project. *Government Information Quarterly*, 30(4), 508-513.
16. Khalid, M. S. B., & Shafiai, S. B. (2015). Flood Disaster Management in Malaysia: An Evaluation of the Effectiveness Flood Delivery System. *International Journal of Social Science and Humanity* 5(4), 398.
17. Khantong, S. (2018). A design of domain ontology to support flood response management and sharing. *Johor Bahru: Faculty of Computing, Universiti Teknologi Malaysia*.
18. Kim, S., and Lee, W. (2014). Global Data Repository Status and Analysis: Based on Korea, China and Japan. *Library Hi Tech* 32(4), 706-722.
19. Kuziemy, C. E. (2017, August). ICT design for collaborative and community driven disaster management. In *CSHI* (pp. 82-87).
20. Leman, A. M., Rahman Salleh, K. A., Babal, M. N. M., Feriyanto, I., Johnso, D., & Hidayah, S. N. (2016). A review of flood catastrophic management in Malaysia. *ARNP Journal of Engineering and Applied Sciences*, 11(14), 8738-8742.
21. Maidin, S. S., Othman, M., & Ahmad, M. N. (2017). Governance of flood disaster framework: A way forward using the framework. *6th International Conference on Computing & Informatics, Sintok: School of Computing*, pp. 650-655.
22. Milinovich, G. J., Magalhães, R. J. S., & Hu, W. (2015). Role of Big Data in the Early Detection of Ebola and Other Emerging Infectious Diseases *The Lancet Global Health*.
23. Rafi, M. M., Aziz, T., & Lodi, S. H. (2018). A comparative study of disaster management information systems. *Online Information Review*, 42(6), 971-988.
24. Santos, L. S., Sicilia, M. A., & Garcia-Barriocanal, E. (2019). Ontology-Based modeling of effect-based knowledge in disaster response. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 15(1), 102-118.
25. Searle, J. R. (1995). *The construction of social reality*. New York: The Free Press.
26. Soden, R. (2017). Data management throughout the national risk assessment process. *UNISDR*, 1-7.
27. Yu, M., Yang, C., & Li, Y. (2018). Big data in natural disaster management: a review. *Geosciences*, 8(5), 165.
28. Yusoff, I. M., Ramli, A., Alkasirah, N. A. M., & Nasir, N. M. (2018). Exploring the managing of flood disaster: A Malaysian perspective. *Geografia-Malaysian Journal of Society and Space*, 14(3).
29. Zakaria, N. H., Ahmad, M. N., Mohd Noor, M. S. A., & Ahmad, M. 2018. Knowledge integration among flood disaster management team: lessons from the Kemaman district. *Journal of Information and Communication Technology*, 17(3), 393-408.

AUTHORS PROFILE



Mohd Ismawira Mohd Ismail is Chief Executive Officer at Integrasi Erat Sdn Bhd. He is currently a PhD candidate under Dr. Mnazir's supervision.



Dr. Mohammad Nazir Ahmad is an associate professor at Institute of Visual Informatics, Universiti Kebangsaan Malaysia (UKM). He received his PhD from University of Queensland (UQ), Brisbane, Australia.



Dr. Nor Hidayati Zakaria is a senior lecturer at Azman Hashim International Business School, UTMKL. She received her PhD degree from Queensland University of Technology (QUT), Brisbane, Australia.



Dr. Razatulshima Ghazali is ICT consultant at MAMPU, Malaysia. She received her PhD from Universiti Teknologi Malaysia (UTM).