Big data analytics nuclear security framework

Maizura Ibrahim1,a, Siti Nurbahyah Hamdan1, Saa’idi Ismail1, Mohd Fauzi Haris1, Mohamad Safuan Sulaiman1, Sufian Norazam Mohamed Aris1, Mohd Hasnor Bin Hasan1, Mohd Dzul Aiman Aslan1, Nur Fatini Abdul Ghani1, and Amy Hamijah Ab Hamid1

1Information Technology & Application Centre, Technical Support Division, Malaysian Nuclear Agency, Bangi, 43000 Kajang, Selangor Darul Ehsan, Malaysia.

a maizura@nm.gov.my

Abstract. Nuclear security is defined as the prevention and detection of, and response to, theft, sabotage, unauthorized access, illegal transfer, or other malicious acts involving nuclear material, other radioactive substances, or their associated facilities. Whereas, big data analytics is denoted as a new way of collecting, analyzing large amounts of data, finding the appropriate patterns to support decision making, hence improve the action taken in solving the problems. The problem with the nuclear security system in most of the nuclear organizations and nuclear facilities in the world today is that the unintegrated data management in nuclear security systems is causing sub-optimal efficiency in the detection of malicious acts involving nuclear/radioactive materials. We argue that a big data analytics framework for nuclear security should be created and utilized to integrate all data in nuclear security systems as well as serving for needs of intelligence especially involving predictive analytics, both within organizations and inter-organization. Therefore, in this paper, we present a conceptual framework of big data analytics for nuclear security. The framework is formulated based on a holistic methodology with the aims to integrate all the data of the nuclear security systems at the organization/facilities level and national level so that the data could be analyzed to derive the appropriate patterns, hence increase the accuracy of decision making that could lead to increased efficiency of detection process and response to the nuclear security event.

1. Introduction

Information technology is the key driver for the development strategy and progress in most areas. The nuclear security area is among the affected areas. Development in information technology contributes to enhancing the quality of security systems, speeding up getting data for security monitoring purposes, effective security management, emerging novel forms of security system, improving communication and interaction between security personnel, and providing access to a wide range of information for further security improvement. The fast evolution of the IT field and its rapid obsolescence stimulate demand for new research, development, and application in these areas. Thus, the evolvement of big data analytics technologies is leading to innovative changes in strategy for nuclear security. A concept for the development of a combination of nuclear security strategy with big data analytics technology could be formulated to provide the solution to problems facing by most of the nuclear organization and nuclear facilities in the world, i.e. a ‘silo’ data management practice for a nuclear security system that is causing a sub-optimal nuclear security system [1].
Based on IAEA Nuclear Security Series, nuclear security could be defined as the prevention and detection of, and response to, theft, sabotage, unauthorized access, illegal transfer, or other malicious acts involving nuclear material, other radioactive substances, or their associated facilities [2]. Thus, any system that is created and implemented to achieve nuclear security, whether it is a computer-based system or a manual system, is regarded as nuclear security systems. Apart from that, big data is denoted as high volume, high velocity, high variety, and high veracity, i.e., the 4Vs of data (information assets) that require new forms of processing, to enable enhanced decision making, insight discovery, and process optimization [3]. Therefore, this paper reveals a new conceptual framework adopting big data analytics technology with various other sub-technological solutions, concepts, and principles as a strategic IT-ecosystem adopting big data analytics for enhancing nuclear security.

2. Related works
This work is in the preceding of two reports published by the World Institute of Nuclear Security in 2015[1][4]. This is also an extended version of our unpublished works that started from 2017 until 2020[5][6][7]. Based on our literature review, there exist several research initiatives in applying big data analytics in the nuclear sector [8] adopting big data for an integrated disaster management system that comprises of flood management and nuclear disaster management. Research on applying big data analytics for the operation and maintenance of nuclear power plant has been conducted by [9][10][11]. It is also found by Roh [12] that adopting the big data analytics approaches specifically using sentiment analysis of social media, to get an insight into the public acceptance of nuclear power in South Korea [12]. However, up to now, we have yet to find evidence that there exists research in big data analytics specifically for nuclear security [13]. Therefore, a new conceptual framework of big data analytics for nuclear security has to be developed from scratch. The fundamental elements of our framework are based on the existing International Legal Framework for Nuclear Security, which we constituted as a non-big data analytics nuclear security framework. The International Legal Nuclear Security Framework has been illustrated by the International Atomic Energy Agency (IAEA) in multiple publications of the IAEA Nuclear Security Series, staggered in four different levels of grouped materials, i.e., the Nuclear Security Fundamentals [14], Recommendations, Implementing Guides, Reference Manuals, and Technical Guidance.

3. The proposed conceptual framework
This conceptual framework is developed in the context of a data management point of view. This concept illustrated how data from the nuclear security system should be managed and integrated so that real-time and autonomous analysis could be carried out to improve or optimize the required situational awareness mentioned in [1] and [4]. In this context, we follow the term ‘Integrated Data Analytics’ introduced by [1] which describes the approach in managing organizations that are led from the top and, where all information and data are used to generate value and support the success of the organization. This framework is also tailored to the statement that security must be viewed and operated as an enterprise-wide activity and be fully integrated into other business processes and objectives [1]. Furthermore, for autonomous analytics, the related data from nuclear facilities/operators must be fed in real-time into the state-level data management system.

Here, we propose a conceptual framework that is developed on two levels. The first level is dedicated to the implementation at the nuclear facilities/operators’ level. The second level is proposed for a state-level implementation by the related state regulatory. However, at this stage, the focus of this framework is only limited to physical security [14][15], personnel security [16][17], nuclear/radioactive materials accounting and control system [18]. Additional elements namely, data visualization, data acquisition, data cleaning, and data exploration are adapted into the framework from the big data analytics concept. The next section will explain the framework that is dedicated to being implemented at the nuclear facilities/operators’ level.
3.1. Level 1: Framework for nuclear facilities/operators

The first level of the proposed framework is depicted in Figure 1. This framework is named the integrated and intelligent nuclear security, I2NS framework. To realize the big data analytics for nuclear security at the state level, the security system that falls into the regime of nuclear security at facilities/operators needed to be fully digitalized. All manual systems should be upgraded into a computerized based system so that the related records and files could be stored electronically in a digitalized form utilizing existing database technology. Depicted in the lowest level of the framework illustrated in Fig. 1 are four components that are critically important for digitalization, i.e. the physical security system, the nuclear/radioactive materials accounting and control system, the personnel trustworthiness evaluation system, and the data visualization component.

Apart from that, all components are linked to the integrated online alert system so that trigger could be initiated from each of the components, whenever the abnormality is detected, as depicted in the second level of the framework. For each of the components, the related back-end databases must be also integrated to form the data pond or data lake, depending on the needed data storage capacity for the respective facility/operator.

![Figure 1. A big data analytics nuclear security framework for facilities/operators.](image)

Figure 2 presents the recommended integrated database architecture that consists of the respected data components of physical security, personnel security, and nuclear/radioactive materials accounting and control. The Apache Hadoop [20] framework is recommended here because instead of relying on expensive, and different systems to store and process data, Hadoop is open-source software that enables distributed parallel processing of huge amounts of data across inexpensive, industry-standard servers that both store and process the data. It is important to state here that before the data could be used to form the data lake/data pond/data warehouse, the data need to be cleaned to ensure the correctness of the data, maintain the high-quality level of the data, and to ensure the data integrity. After that, the data must be explored to develop the analysis model that could lead to the development of data visualization and analytics product, that is usually displayed as a dashboard system.

In this framework, we have identified three types of data that must be dealt with to develop the big data analytics nuclear security system at the facilities/operators’ level, which is, structured data, semi-structured data, and unstructured data. Structured data is data that is extracted from databases that may consist of a multi-platform database management system such as Microsoft SQL Server, MySQL, MariaDB, Oracle, or MongoDB. The semi-structured data may come from the radiation portal, sensors, or text files related to the nuclear security system. Lastly, unstructured data is the most difficult data type to be handled. This kind of data may come from the CCTV monitoring system, or surveillance system, which consists of images and video stream frames.
As depicted in Fig. 2, the personnel trustworthiness evaluation is included as a component that utilized the sentiment analysis of the media social approach. The result of the sentiment analysis is then injected into the personnel profile databases for the evaluation of personnel trustworthiness. The trustworthiness evaluation is important for the detection of insider threats.

![Figure 2](image_url)

**Figure 2.** The integrated database architecture.

### 3.2. Level 2: State-level framework

Figure 3 illustrates the state-level framework. The elements in the framework are designated based on the nuclear security elements stated in the International Legal Nuclear Security framework and the need to perform autonomous analysis of the data for the realization of the enhancement situational awareness. It is recommended that the state-designated regulatory bodies that are appointed to safeguards nuclear/radioactive materials to take charge of owning the respective system based on this framework.

In this framework, we had specifically identified the external data that is needed to be acquired from external parties such as nuclear facilities/operators and GIS/GPS data providers. The data that is required from the related regulatory bodies is also stated in the framework. Once the data acquisition is done, data cleansing and exploration needed to be performed to develop the state-level analysis model that could result in the finding of a data product that could give an insight of what, where, when, and how the improvement shall be made with regards to the state-level nuclear security.

![Figure 3](image_url)

**Figure 3.** State-level big data analytics nuclear security framework.
The output could be an improvement of the Act, guidelines, or other related regulations tailored to the security, safety, and safeguards of the nuclear/radioactive materials which will be beneficial to the public, nuclear organizations, and other stakeholders.

4. Conclusion and future works
Nuclear security is the process of prevention and detection of, and response to unauthorized removal, sabotage, unauthorized access, illegal transfer, or other malicious acts involving nuclear or radiological material or their associated facilities. The possibility of nuclear material or other radioactive material that could be used for criminal purposes or intentionally used in an unauthorized manner cannot be ruled out in the current global situation. Currently, most information and data in nuclear security applications are managed in silos both within and between organizations which resulted in many current data management systems for the nuclear security system that are suboptimal hence cannot provide support to improve security decision and control.

In this paper, we had proposed a conceptual big data analytics framework for nuclear security. The framework is developed in the context of a data management point of view. It consists of two levels. The first level is dedicated to the implementation of nuclear facilities/operators. The second level is for the state-level implementation. This framework has yet to be evaluated. Therefore, our future works are to evaluate this framework by applying the empirical methodology. We are currently in the process of developing survey questionnaires. We had reviewed some related techniques for the evaluation process and choose to implement the Rasch Measurement Model for the evaluation process.

References
[1] World Institute of Nuclear Security, WINS 2015 Data analytics for nuclear security, how real-time integrated data management could support nuclear security WINS Special Report Series, (Vienna, Austria).
[2] International Atomic Energy Agency IAEA 2008 Security in transport of radioactive material, implementing guide. IAEA Nuclear Security Series.9 (Vienna, Austria).
[3] Beyer M A and Laney D 2012 The importance of big data: A definition. (Stamford: CT: Gartner).
[4] World Institute of Nuclear Security WINS, 2015 Big data in motion, how real-time integrated data management could support nuclear security WINS Special Report Series (Vienna, Austria).
[5] Ibrahim M, Ismail S, Haris M F, Hamdan S N, Sulaiman M S, Hamid A H A, Aris S N M, Aslan M D A, Ghani N F A and Hassan H 2017 A proposal on adaptation of big data analytics framework in nuclear security toward development of integrated, real-time, and intelligent nuclear security system Nuclear R & D Seminar 2017 (Selangor, Malaysia).
[6] Ibrahim M, Ismail S, Haris M F, Hamdan S N, Sulaiman M S, Hamid A H A, Aris S N M, Aslan M D A, Ghani N F A and Hassan H 2018 The integrated intelligent nuclear security (I2NS) system: A progress report for 2018 Technical Report No.Nuklearmalaysia/L/2018/173 (Kajang, Selangor, Malaysia: Malaysian Nuclear Agency).
[7] Ibrahim M, Ismail S, Haris M F, Hamdan S N, Sulaiman M S, Hamid A H A, Aris S N M, Aslan M D A, Ghani N F A and Hassan H 2020 Malaysian Nuclear Agency integrated and intelligent nuclear security (I2NS) system: A progress report 2019 Technical Report No.Nuklearmalaysia/L/2020/10 (Kajang, Selangor, Malaysia: Malaysian Nuclear Agency).
[8] Pu C and Kitsuregawa M 2013 Big data and disaster management a report from the JST/NSF joint workshop Technical Report No. GIT-CERCS-13-09; (USA: Georgia Institute of Technology, CERCS).
[9] Jae-Min C, Junguk S, Choong-Sub Y 2015 A Review on applicability of big data technology in nuclear power plant: focused on O&M phases Trans. of the Korean Nucl. Soc. Spring Meeting, May 7-8 (Jeju, Korea).
[10] Moon-Ghu P, Seong-Ki K, Jangbom C, Ji-Hun K 2017 Development of a big data framework
for operation support of nuclear power plants Proc. of the KNS 2017 Spring Meeting.

[11] Liu K 2019 Big data analytics solutions to improve nuclear power plant efficiency: online monitoring, visualization, prognosis, and maintenance decision making Adv. Sensors and Instr. Ann. Webinar, Nov. 6, 2019 (University of Wisconsin-Madison, USA).

[12] Roh S 2017 Big data analysis of public acceptance of nuclear power in Korea J. of Nucl. Eng. and Technol. 49 850-854.

[13] Ibrahim M, Ismail S, Haris M F, Hamdan S N, Hamid A H A, Aris S N M, Aslan M D A, Ghani N F A and Hassan H 2020 big data analytics for nuclear security: a literature review Nuclear R & D Seminar 2020, 16 -19 Nov. 2020 (Kajang, Selangor, Malaysia).

[14] International Atomic Energy Agency IAEA 2013 Objective and essential elements of a state’s nuclear security regime, nuclear security fundamentals. IAEA Nuclear Security Series 20 (Vienna, Austria).

[15] International Atomic Energy Agency IAEA 2013 nuclear security systems and measures for the detection of nuclear and other radioactive material out of regulatory control, implementing guide IAEA Nuclear Security Series 21 (Vienna, Austria).

[16] International Atomic Energy Agency IAEA 2015 Risk-informed approach for nuclear security measures for nuclear and other radioactive material out of regulatory control, implementing guide. IAEA Nuclear Security Series 24-G (Vienna, Austria).

[17] International Atomic Energy Agency IAEA 2008 Preventive and protective measures against insider threats, implementing guide IAEA Nuclear Security Series 8 (Vienna, Austria).

[18] International Atomic Energy Agency IAEA 2020 Preventive and protective measures against insider threats, implementing guide. IAEA Nuclear Security Series (Rev. 1) 8-G (Vienna, Austria).

[19] International Atomic Energy Agency IAEA 2015 Use of nuclear material accounting and control for nuclear security purposes at facilities, implementing guide IAEA Nuclear Security Series 25-G (Vienna, Austria).

[20] https://hadoop.apache.org accessed on 13th November 2020.

Acknowledgments

This study is supported by the International Atomic Energy Agency (IAEA) research grant under the IAEA Coordinated Research Project CRP J02011: Improving the Security of Radioactive Material throughout its Lifecycle, Associated Facilities, and Associated Activities (Contract No. 22165).