The Impact of Big Data on Risk Management and Methods to Reduce Risk

CSCI 6444

SUMMER 2021

Kahang Ngau

G26462030

Jul 19th, 2021

1. Introduction

The digital era has grown rapidly nowadays. People use online services heavily and widely in different aspects of their lives. Meanwhile, the use of online services has also become a key for organizations in such a digital era as well. Online customers have grown tremendously and are now becoming an increasingly large segment of today's consumers. Even if you are the seller but do not sell goods or services online directly, your online platforms will be the place where your customers often begin their exploration to seek out your services. And this can bring up the question of how secure and reliable your online services are. According to numerous cyber security reports, the number of cyber threats is increasing rapidly from 23,680,646 in 2008 to 5,188,740,554 in 2013 (Jasiul, 2014). As we all know, every organization faces unexpected risks. And we should all know that sometimes the risk is inevitable on the path to achieving success and it is hard to predict. It will be ideal for an organization to eliminate all the risks, but it is almost impossible to accomplish them. These threats, or risks, can come from a big variety of sources. These sources can come from financial uncertainty, strategic management errors, cybersecurity issues, accidents, and natural disasters. As a result, risk management strategies are critical to alleviating information technology security threats and data-related risks. It becomes the company's or organization's top priority since it helps them to identify and control risks to its digital assets. And all these assets can include corporate databases, customers' personal information, intellectual property, etc. If all these elements are protected safely by the organization, harmful events that can cause the organization to lose in time and money can be reduced or well informed ahead of time. This is an interesting and important topic to discuss because risks or threats are everything and it relates to every aspect of a company or organization. It has a great influence on the organization's security issue and can impact the organization in the long run. By knowing well where the places that risks or threats can occur, organizations can formulate risk management plans to establish procedures to avoid these potential risks and minimize their impact when they occur with the result.

It is well-identified that information technology plays a critical role in the digital era in many businesses. And it will be helpful and important for anyone who owns or manages a business that makes use of IT to effectively identify risks to the IT systems and data information, by developing responses and rescue plans in the event of an IT crisis taking place. The utilization of high technology not only helps the organization to manage risk by developing control focus, but also it remains as one of the important tools to enhance the IT systems both commercial and communal. Risk management has played a major role in the field of business and it was realized that many risks can be prevented through loss-prevention and control systems. And many of these systems are developed by using advanced computational models and techniques. In this section, various tools and techniques that are implemented in IT systems to reduce risk will be introduced and discussed.

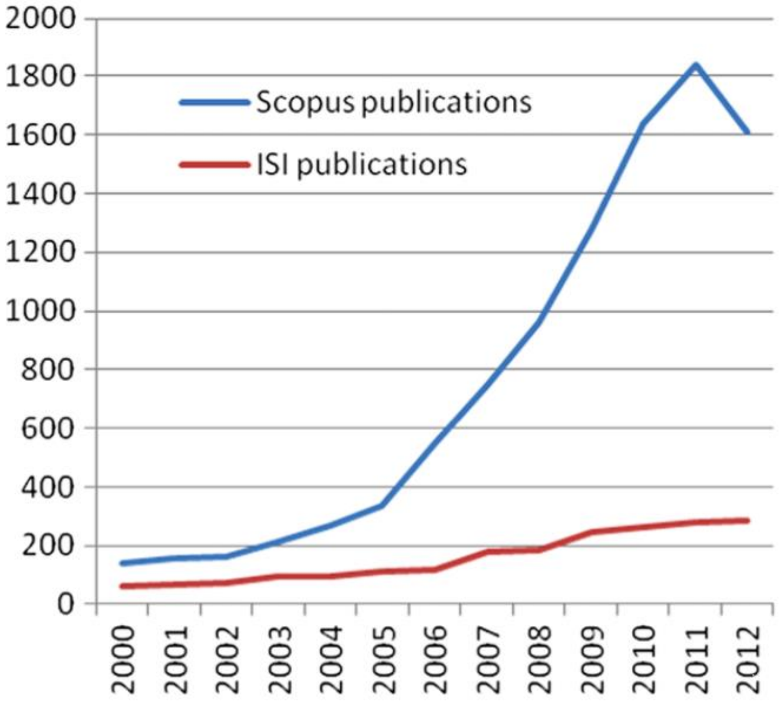


Fig. 1. Charts for publications on ‘‘enterprise risk’’ since 2000, from Wu’s paper published from Preface / Mathematical and Computer Modelling 58 (2013) 1581–1587.

The graph from Fig. 2 presents the number of publications on the topic of enterprise risk since 2000. The result was created by filtering the publication by the keyword ‘enterprise risk’ in both Scopus and ISI Web of Science. All these publications are verified by the third party that certain levels of techniques are performed when managing various risks in the enterprise. Both curves from the graph show that there is a trend for published articles in enterprise risk management to have increased during the past decades (Wu, 2013). In general, IT risks can include failure in hardware and software, spam information, virus attack, human error, and natural disasters like fires or floods. Hardware and software failure occurs when the organization has a power outage or data corruption. Spam is an unsolicited email that seeks to fool people into revealing personal details or buying fraudulent goods. A virus attack refers to computer code that can copy itself and spread from one computer to another to disrupt computer operation. The human error refers to incorrect data processing and data disposal which leads to IT system concerns. If the organization or the business relies on the information technology system in which computer networks are the key operating activities, they should pay more attention to the scope and the nature of these IT systems. As previously mentioned, those risks or threats are identified as general risks in an IT system. There is another kind of risk or threat that is identified as criminal IT threats. Criminal IT threats include hackers, fraud, passwords theft, denial-of-service, security breaches, and staff dishonesty. Hackers describe people who illegally break into an organization's computer systems. Fraud occurs when someone uses a computer to alter data for illegal benefit. Password theft often refers to the target of malicious hackers. Denial-of-service is an online attack that prevents website access for authorized users. Security breaches include physical break-ins as well as online intrusion. Staff dishonesty refers to someone's behavior related to stealing data or sensitive information. After learning all the kinds of risks or threats that can occur in an IT system, organizations need to develop effective risk management and assessment plans to minimize the impact of all kinds of these risks on the organization in the long run.

2. Background

In the digital era, big data is evolving the practice of risk management and assessment. The use of big data has greatly enhanced its importance in the relationship between business and risk management. With the growing need for digital services and information acquisition online, companies can retrieve data about potential customers and users' behavior. After gathering all these kinds of information, the organization conducts data analysis to improve its risk management practices in a more effective and advanced way. One of the biggest influences that big data has on risk management is to identify and provide emerging trends and risk factors by learning from past data. For instance, companies make good use of big data by identifying emerging and existing trends in their current customers. Through conducting different statistical analyses, companies can create new business plans in detail aimed to match with the shifts that are newly detected in user behavior early. This can potentially alleviate the risk of changing the direction of a business for the company. Moreover, the conducted statistical analysis can also be useful for the company to identify important factors that contribute to customer defection. After knowing some of the important factors, companies can perform machine learning in models to train the dataset and use the model to predict results based on the input factors. The prediction results can also be an important reference for the company to make adjustments for their business plans. Although the digital era has made some risks or threats appear that did not exist a couple of decades ago, it has made and created more advanced and effective solutions available for people to manage different kinds of risks. For instance, companies in financial industries that deal with sensitive information like personal information or financial factors can make good use of big data to potentially identify frauds and threats by analyzing the risk or threat factors from the database. It allows organizations to detect unusual behavior and discrepancies from the data by performing a highly streamlined and filtering process in detail. To complete the previously mentioned process, advanced modeling techniques including coding and model building are needed to form an automation process. The automation process not only increases the accuracy which greatly improves the quality of the results but also saves a lot of time in human labor by reducing repetitive work. No more wasting hours of manpower can potentially reduce human error which is one of the general IT risks introduced in this paper earlier.

According to the statistics that are provided by big data, companies that deal with risks in credit card fraud, market risks, and asset-liability are reported more in need of risk management and assessment than those financial institutions (Cole, 2015). This is interesting to notice that not only financial institutions need risk management heavily, but companies that deal with any kind of data would also require proficiency in risk management to protect their information safely and improve their customers’ satisfaction. There are many benefits that big data can provide but along with these advantages comes risk. At the time when big data was manipulated to reduce risk in IT systems, new risks are shown in the area of data security and data rights. Since previously mentioned that companies use a great amount of big data to analyze and make predictions, the big challenge of storing all these kinds of data arises. It will be costly and not economical to store these huge amounts of data in companies’ available hardware and software. However, data safety issues are concerning when the companies’ data is outsourced to other cloud computing services while making the company available to focus on developing its own core business. Again, nothing is perfect just as relying on others to store a company’s data can potentially put the company at risk by losing its control over its data. It will be key for companies and organizations to find a balance to effectively make good use of the big data to minimize the potential risk in their IT systems. This term paper intends to provide an overview of the impact of big data in IT systems and to analyze different methods that can be used to reduce risk in different industries.

3. Benefit and Risk of Big Data

Lots of great benefits and advantages of using big data are introduced earlier in the paper. However, the practice of big data also has a great positive impact in the field of medical research. According to Cole's paper, scientists and medical professionals alike are capable of studying more patients at a higher degree of speed and accuracy which greatly increases their research efforts (Cole, 2015). researchers in the medical field utilize big data to speed up their research process. One example would be to use big data to make digital maps globally for infectious diseases. It is very helpful for researchers to understand the disease outbreak range geographically ahead of time and make practical medical plans accordingly. Furthermore, big data can also be used to determine the relationship between the disease outbreak and the environmental factors. After conducting statistical analysis, researchers can see the distribution of the geographical regions that have the most cases of the disease occurring within a particular time range. The environmental factors or elements of the region should be identified and recorded so that researchers can use these environmental factors to make predictions about where there will possibly be another disease outbreak region. However, the process in the medical research field is moving slowly due to the limitation in human resources which delays often happen. Although it is a great direction for the medical research field to analyze the relationship between disease and environmental factors using big data, sometimes it will still be difficult to create the entire database only based on the occurrence of the diseases. Another great way of using big data in the medical field is to provide better or more accurate treatments to patients based on similar conditions that can be found from the past in the database. Instead of trying different kinds of treatment and not knowing what the outcome will be, doctors now can use big data to similar patients in the past based on the patient's genetic makeups, diagnostics, lab reports, and various other similarities (Cole, 2015). In this way, physicians can use the searched results as a reference to make adjustments on whether the treatment should work or not. This can greatly improve the success rate for physicians to treat patients based on another patient's past treatment. Big data has become an advanced tool in the medical field which empowers medical research and medical treatments.

As we already know, big data can bring along with it risk at the time when people use it. The data privacy issue is one of the biggest concerns in the use of big data, especially in the medical field. Since lots of sensitive data like patients' personal information and patient's treatment records are stored and gathered in systems, patients are afraid of the safety of their personal information being exposed. There is also a rising concern on the use of open-source systems on storing data, for example, NoSQL. The concerns stated the problem that many of these open-source systems fail to maintain a certain quality level of security which brings them in system risks. Due to this shortcoming, people who manage with these systems should drive and interpret the data in the correct direction. Problems like inaccurate results, misleading trends, incorrect cost analysis, and credit crisis may appear when the data is driven in the wrong direction. Although there is still a long way to go for people to precisely determine whether open-source tools should be used or not, the goal for the end-user is to have all aspects of an accurate and secure system to store their information and data. In the field of medical practice, big data has now become an advanced modern medicine that can track disease and provide reliable patient care while the data right and security issues remain a challenge in this field. The internal and external security issues are also urged to be fixed, otherwise, data and information can be manipulated and mislead people that causing bad consequences.

4. Role of Risk Management in IT System

For any business, risk assessment and management are some of the best ways to prepare for eventualities that can occur or are in the way of process and growth. Risk management is an important process for a business because it enhances the business's ability to adequately identify and deal with potential risks. It is a great way for the organization to work closely with the IT manager to make a balance on operational and economic costs for achieving specific missions or goals. according to Tohidi's paper, administrators of each organization unit must be assured that the organization has the needed ability to achieve a certain mission. They can provide the best conditions for encountering missions with real-world behavior with the determination of security abilities (Tohidi, 2010). Risk management is the key element for helping managers or executives of a company to control and make decisions on the maintenance of IT systems. One of the important processes in risk management is to use and fully integrate with the system development of the life cycle. It involves five steps which are the beginning stage, development stage, operation stage, protection stage, and administration stage. The first three phases in the SDLC help the IT system to identify risk, define risk, and support the evaluation of the system operation. The protection phase involves adding software and making changes in the process, organization policies, and customs which improves the system's operational functions. How risk management can be used in this phase will be renewing credit of the periodic systems when there are essential changes in the production-operation environment of the system occur. The final phase of administration includes actions associated with activation, filing, rejection, or destruction of information in the system. And the risk management in this phase will be implemented to ensure that appropriate hardware and software are installed and licensed within the system.

After integrating the System Development of Life Cycle, the first process that needs to step in is to establish risk assessment plans to focus on the attention to the risk link with the IT system through the SDLC. This step can effectively determine the danger of the minimal risk, and this process will be helpful for the organization to identify any useful technical controls that can be performed to reduce or limit the risk or threat during the process of risk reduction in the system. This is also defined as the risk estimation method which is one of the important steps during the risk reduction process. To understand clearly the vulnerability and potential threats to the system as well as to determine the possibility of a failure adverse events are essential to the IT system. According to Tohidi's opinion, there are steps in the risk estimation process and which are, system characterization, threats, and vulnerabilities identification, analysis of the controls, probability determination, effect analysis, risk determination, control purchase order, and documented results (Tohidi, 2010). The first step is system characterization which involves defining the action scope, the boundaries of a valid application, and resources and information of IT system boundaries. And the ideal output for this step is to provide a clear and outstanding blueprint of the IT system environment and boundaries. The assessment of the risk in the IT system will also be able to be executed when the field for risk assessment is imposed. After characterizing the system, the second critical step will be too well identify the threats or risks in the system. Checking and considering threats sources, potential vulnerabilities, and the existing controls are the key actions for determining threats or risks in an IT system. And it will be desirable to list all existing threat sources or potential threat sources so that we can focus on these targets to take action. Once the list of threats is generated, the IT system should be able to conduct analysis on these threats which includes analyzing the vulnerabilities of the system environment. The purpose of this action is to specify damages coming from threatening sources. The development of the system vulnerabilities list which is deficiencies and weaknesses is often the target for threatening sources to attack from. The next step is to analyze and classify the controls that are applied to the organization. Creating a list of current existing or designed controls for the system can minimize or limit the possibility of threats coming from some particularly vulnerable systems. Moving forward, there is a need to determine the rate of the likelihood of the damage being achieved from the threatening sources. And it is important to understand well what is the motivation and ability of the threats, the nature of the vulnerability, and some of the existence and impact they have on the current controls. It will be wise to place the threatening source in high, low, or medium-level so that each level of risk should be matched with corresponding solutions. Besides rating on the damage from risks, determining the size of the effect of purpose by analyzing the level of the effect associated with the organization’s information assets. It provides a great understanding of the qualitative or quantitative sensitivity and criticality of these assets and how they should be prioritized. And finally, the step for controlling ordering takes place to further reduce or eliminate the risks or threats that are identified in the system. The purpose of this work is to reduce the level of risk of an IT system while keeping data security at an acceptable level. Given that the system’s data sensitivity can be based on the level of system protection requirement, certain factors in the controlling ordering process should be considered, including the legislation and code, organized policy, operational effect, and safety and reliability (Tohidi, 2010). These are the side solutions to mitigate risk in the risk estimation process. All these steps in risk estimation should be recorded and reported as a management report to help inform the organization’s managers or executives to make any business or operation decisions regarding the risks. The report should be reviewed thoroughly by the risk management team.

The second process of risk management after risk estimation is to perform risk reduction. In this process, risk reduction controls the prioritization and evaluation of certain applications, since it is nearly impossible to eliminate all risks or threats in an IT system. It will be ideal for the organization to use the least cost but the most effective approach to minimize risk to an acceptable level. Making classification of controls over the IT system can be one of the good approaches for reducing risks. And this process has three different focuses according to Tohidi's paper and which are, technical security controls, management security controls, and operational security controls. Technical security control involves establishing protection by building a set of security packages on the organization's hardware, software, and firmware. It effectively helps the system to fight against all existing risks or threats. Management security control has a certain degree of attachment with technical security controls. It follows rules, guidelines, and standards to protect information safety throughout the operational procedures of the organization to achieve goals and missions. Operation security control sets up security standards of the organization to ensure security procedures monitoring are inappropriate to use on IT resources and assets. The use of these three different kinds of controls requires an organization's staff to be trained appropriately so that the goal of risk reduction can be achieved. After identifying all possible controls from all these three perspectives, the organizations can evaluate their feasibility to allocate their resources efficiently and enforce effective controls on their systems. The cost of analysis can also be performed at this stage so that organizations can have a better understanding of each proposed control to determine which one of the controls is appropriate in certain conditions. The purpose of the cost analysis can be beneficial to the organization's research team to better determine whether the cost of applying the controls in the IT system to reduce risks or threats is cost-effective or not. Managers or executives from the organization can evaluate and adjust the business plans according to the performance of the different control approaches. Communication networks are extremely essential in most organizations nowadays due to the constant growth in the need for collaboration in the digital era. The organization's applicable software is constantly updating which indicates that new levels of risks and previously reduced risks may occur again. This has become a potential concern as the organization's personnel and security policy are altered. Therefore, risk estimation and reduction processes should be repeated at least once in three years for organizations so that they can integrate with the SDLC to better support their business and mission goals in the long run. Understanding the risk management process is essential as it plays an important role in securing the safety of the IT systems in any organization. And this goal can be achieved by getting support from the board, participation of managers, executives, and trained staff from the organization.

5. Mathematical Method to Reduce Risk

In this section, a real-work problem will be discussed applying the mathematical method and model to reduce risks and uncertainties by selecting risk response strategies. As we all know that risks and uncertainties always exist in various IT systems, the failure in accurately identifying the risks and uncertainties can greatly increase the total cost and may also cause essential social and environmental damage (Cheraghi, 2017). The purpose of the study from Cheraghi's paper is to develop a mathematical programming model to select risk response strategies for construction projects. And the strategies that are mentioned in the study can be applied to many other similar analytics projects since all-risk response strategies need to be identified and concluded. The mathematical programming model is presented with the time, cost, and quality that are measured to obtain an optimal risk response strategy for the project. The first step in this method is to perform risk identification where the potential risks are identified and appropriate strategies should be conducted in responding to risks. Failure Mode and Effects Analysis (FMEA) approach is developed and it is the most effective method to evaluate risks for complex projects. In the past couple of decades, responding to risks is one of the vital steps for the most proposed models by different scholars. Analysis checklists, SQOT approach, Delphi techniques, and multi-criteria decision-making techniques such as AHP are some examples of their techniques used in this field (Cheraghi, 2017). Therefore, the FMEA method is very useful in evaluating and identifying risks in various projects. There will also be an optimal model for selecting the risk response strategies afterward.

Before the mathematical programming model is developed for risk response selection, a work breakdown structure analysis should be performed to integrate the comprehensive project management with other risk management sub-systems. It is one of the important steps that need to be completed before developing the optimal model. In this approach, several measurements are chosen to meet the system constraints and optimize the objective function. Generally, binary variables are defined in an optimal model to choose the response stratified to the risk of a complex project. The development of an integer programming model is to maximize the effects of the responses to the risks based on the variables like budget, schedule, and quality. The object function of the models includes a set of strategies that will be used to enhance the responses to the estimated risks. After that, risk response strategies are identified and determined after recognizing the risks through applying the FMEA methods. Just to keep in mind that project risks are relatively complicated and they vary within different projects. Setting up assumptions can be an effective approach for simplifying the analysis of the project.

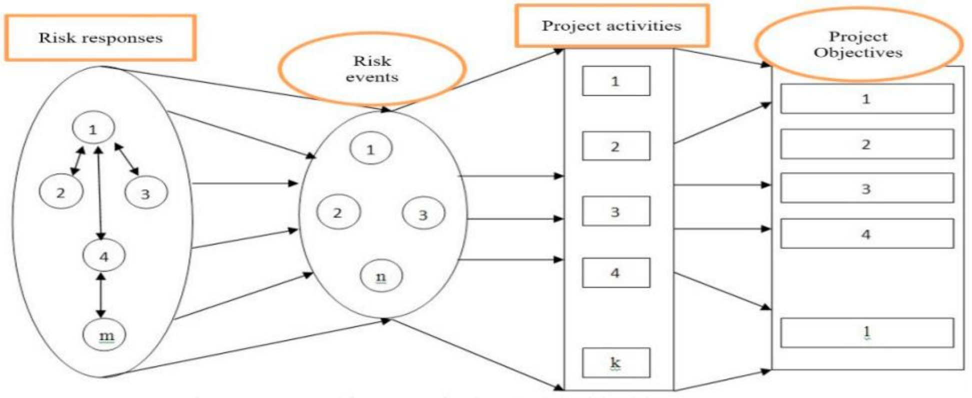


Fig. 2. A framework for selecting the project risk responses from the Cheraghi paper, 2017.

The above framework displayed the risk response strategies model from the case study on Cheraghi's paper. It depicts the model framework with the relationship between the answers provided by the related risks to the independent individuals who answer. The four different groups of design-build activities which cooperate closely with the system's technical and engineering services are well categorized on the framework. Firstly, project risks and critical project activities are arranged and prioritized by the FMRA method. The proposed mathematical programming model with the use in GAMS software can determine all identified risk response strategies. The conceptual model will be helpful to develop so that it can evaluate and select the response to the project risks which have an association with work breakdown structure, risk events, and alleviate the risk and effects to each other. After all, the optimal strategies for responding to the project risk will be chosen from the mathematical model, although there is a cost of risk occurrence and the quality loss of each activity caused by each risk in the model. According to the opinion in Cheraghi's paper, lots of the other studies mainly focused on the cost of risk in human resources, attitudes, and experience related to factors and their relations to binary variables, while this paper considers the Work Breakdown Structure as one of the most important parts for risk identification and assessment (Cheraghi, 2017). In short, it is reasonable to consider using multiple objectives to solve the model when the dataset or the size of the project is relatively small. When dealing with the large size of the project, using the metaheuristic algorithm can be one approach due to its complex characteristics.

6. Computational Simulation and Risk Analysis

After knowing one of the mathematical programming methods for risk analysis, it will be interesting to explore in-depth with other computerized tools in the field of risk management and analysis for IT systems. A new method of risk propagation among associated elements based on the use of colored Petri nets was introduced in Szpyrka's paper. The proposed method demonstrates the model relations between nodes forming the network structure and it takes into account the bidirectional relations between components as well as relations between isomorphic, symmetrical components in various branches of the network (Szpyrka, 2017). This powerful method is intended to assess critical infrastructure risk and it is capable of being adapted with propagation models in various systems. It is demonstrated as an approach to evaluate risks for cyber-attacks in IT systems in correspondence with Petri nets. Eventually, the evaluation of the risk from cyber-attacks can be estimated for a particular node in the model, as well as all other related nodes such as hierarchy relations of components and isomorphism of elements. The aim of this method is not only to help IT systems to improve cyber situational awareness in the fields of system vulnerabilities, threats, incidents, and attacks but also to support decision-makers to react and respond to emergencies in case of cyber-attacks. Based on the characteristics of Petri nets, algorithms are developed in the method to assess the impact of the risk one node has on the other related nodes. And it also consists of the hierarchy and similarity of the nodes vulnerable to the same threat in the network system. In other words, this method not only creates models with risk propagation to reduce risk in the hierarchical structure, but it also considers the bidirectional hierarchical relations between elements in various branches of the tree. This feature provides the ability for the method to model the risk propagation between isomorphic components. For instance, if there is one asset in some branch of a tree that is affected by a threat, it will be desirable for the system to detect that this threat may influence symmetrical or identical assets in other different branches vulnerable to the same threat. Therefore, people can evaluate how the security of one component of the IT system affects cooperating nodes and the whole network as well. And for cases of cyber-attacks, various systems can use this method to conduct behavior modeling to verify additional security adoption in protecting the system infrastructure or particular node as well. The IT managers or administrators of the system can also monitor the status of the risk and inform stakeholders about the negative impact and possible adjustment in the use of the risk propagation method. This method further plays a critical role for decision-makers to take proper actions in protecting the system infrastructure.

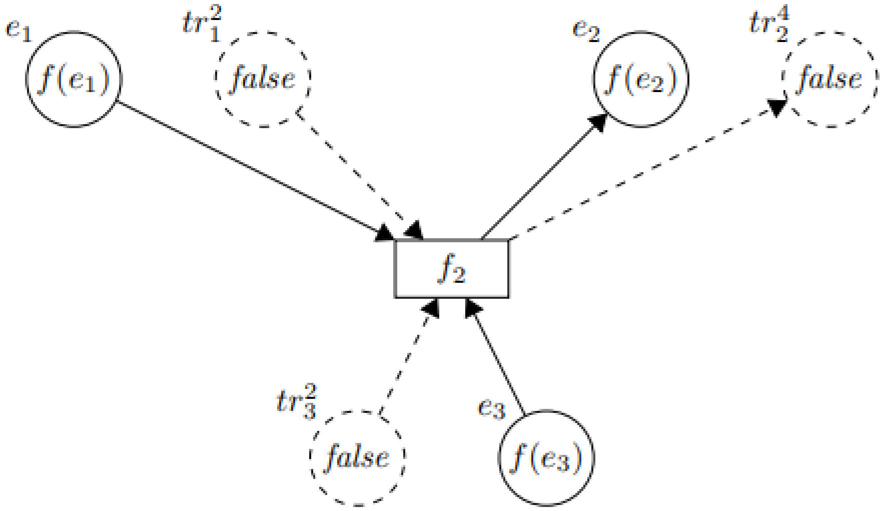


Fig. 3. Part of the propagation net for the considered system represents a function that is used in the method from Szpyrka's paper, in Symmetry 2017, 9, 32.

When speaking of Petri nets, it is one of the powerful mathematical modeling languages in distributed systems and it is widely used in the field of cybersecurity to help systems reduce risk. Malware is one of the most common cyber-attacks that target end-users and IT systems by causing malicious software (Jasiul, 2014). There are ways that Malware can harm the IT system so the detection process should be focused on analyzing behavior models which can support the selection process of cyber-attacks and facilitate the application of countermeasures. The Malware detection process takes a set of suspicious events from the process' hooking engine. This engine filters single malicious incidents among hundreds of thousands of regular ones based on the colored Petri nets Malware models. The CP-nets model has a hierarchical structure and it equally describes the state and actions of the modeled system. As you can see the primary module is presented in Fig 4, where on the left side is a column of places storing tokens for particular assets that may be affected by Malware. The second column consists of substitute transactions that are related to the acquisition process. The third column represents particular assets that are affected by the Malware activated in the monitored system. And finally, all these observed symptoms are processed by the verified transaction which examines if the system is infected by a type of Malware. This model also has the advantage of allowing one to easily update the symptoms of a new attack by changing the initial mark of places. Therefore, this approach for modeling Malware behavior is developed based on the use of colored Petri nets which can effectively detect and prevent advanced persistent threats in IT systems.

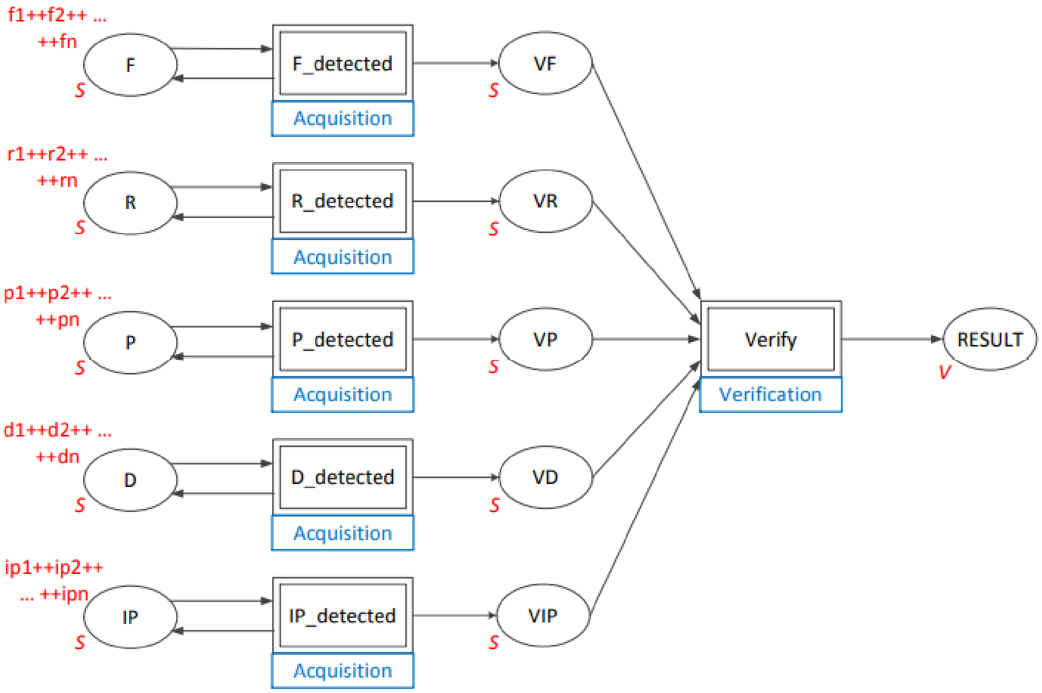
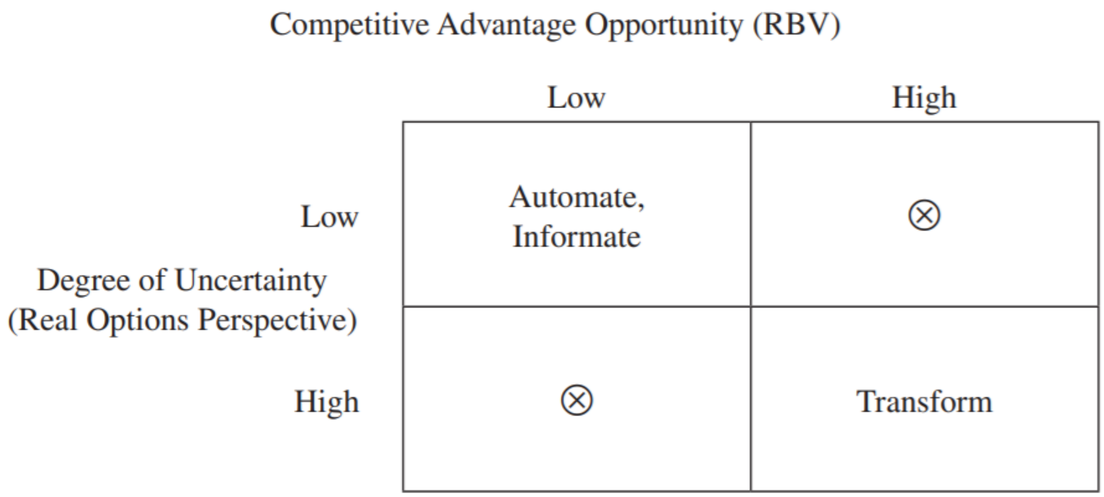


Fig. 4. The primary module of the Colored Petri nets model of the system to detect Malware, from Jasiul's paper, in Entropy 2014, 16, 6602-6623.

One other useful computational simulation method that is introduced from Wu's paper about reducing enterprise risks is to develop optimization tools. The existence of uncertainties and risks in an IT system may make it difficult for optimization tools to obtain the required assumptions. However, some models are created to optimize the engineering system and to minimize potential system risk by considering the increasing probability of failure in the mechanical system due to the aging issue. This approach is very practical because the aging problem exists in any IT system and it will be ideal to create computational programming models to conduct analysis. By using machine learning techniques and utilizing the big database for the past performance of machines, we can conduct a feature selection process to find the key variables to be used in the machine learning model. Before training the data set, the dataset will first be separated into a training set and a testing set with a 7:3 ratio of the original dataset. The training set of the data will be trained by the machine learning models for all of the selected features or variables related to the aging issue of the mechanical equipment. Then by comparing the Root Mean Square Error (RMSE) and R-squared of the results from the models, it will be wise to choose the model with lower RMSE and higher R-square value. The RMSE particularly describes the model performance by measuring the difference between the values that are predicted by models or an estimator and the actual values that are observed. The R-squared value, on the other hand, describes the good of fitness of the models. In other words, it describes the proportion of the variances in the dependent variable that is predicted from the independent variables. After choosing the model with the lowest RMSE and higher R-squared value, the selected model will then be used as the final model to make predictions on the possibility of failure in mechanical equipment due to aging. This approach is essential in the process of reducing risk for systems because it gives people a great understanding on what are the influencers or independent variables that have the greatest impact on the failure of machines. Managers or executives of an organization can react quickly to make decisions and to develop proactive plans for improving the machine failure problem of their system due to the results from the models.

Computational programming tools and investment plans are widely used in the decision-making process in the field of information systems. Lots of these tools or applications involve decision-making and industry modeling under uncertainties and risks. The paper written by Otim discusses a recent analysis especially addressing the evaluation of value and risk in information technology investments. These investments include a complex set of shareholders, leading to the need to consider organizational policies (Otim, 2012). From the paper, different types of IT investments are analyzed since they have a great impact on downsizing or minimizing risk in the IT system. IT investments include transformational and informational which have the capacity in leading strategic IT investment in industries that can potentially meet the goal of risk reduction. For the IT investments that involve computational programming which creates automation in business functions, the reduction in downside risk can be performed through investing in parity with industry participants. In other words, it is important to make an analysis emphasizing the context of IT investment since many investments only focus on short-term advantages. There is a need for organizations to clarify strategies so that the likelihood that the company or firm will underperform relative to its competitors can be limited. From the opinion that Otim states, firms or organizations which have developed transitional IT strategic roles tend to have positive and abnormal returns. And the results support the value of considering the strategic transformation IT role into conditions under which IT investments are likely to produce similar results. It is clear to notice that there is a relationship between the reduction in downside risk and the announcement of strategic IT investments. And it is important to know that if the firms or organizations lead similar transformative IT investments, they have a higher chance to successfully have a reduction in downside risk in their systems. One important concept to help industries to reduce downside risk is to develop the method of resources-based view (RBV). This method indicates that the rarer and more valuable the resources are, the more competitive advantages they have. So, it is suggested that if the investment is being conducted earlier, it will be realized to have first-mover advantages and will be very crucial for IT investments to build network effects.

Another key method about the IT strategic investment roles is the real options method. It emphasizes the importance of managerial flexibility which derives benefits from the method of analyzing how uncertainties can be resolved. Therefore, the timing of investment is very important as well as the ability to perform organizational learning in the resolution of certainty and realization of the value of real options. The below graph demonstrates the framework of the RBV method based on the uncertainty-expected payoff matrix. The framework has the background knowledge that managers or executives of an organization are making the balance of the commercial failure and the opportunity for the gains such as competitive advantages. The RBV for the organization is suggested to develop sustainable competitive advantages by the effective deployment of resources and the guidance for investment decision-making. The real option, on the other hand, indicates that investment decisions are made by the need to reduce and minimize the threats or risks of commercial failure because of the investment uncertainties.



		Competitive Advantage Opportunity (RBV)	
		Low	High
Degree of Uncertainty (Real Options Perspective)	Low	Automate, Informate	⊗
	High	⊗	Transform

Fig. 5. Typology of Strategic IT Investments from the methods of Competitive Advantages Opportunity in RBV and the real option, from Otim's paper in Journal of Management Information Systems, July 2012.

After knowing the important methods and concepts in both RBV and the real option, data mining has also become a very popular approach and technique to apply statistical and artificial intelligence tools to analyze large sets of data sets. The applied data mining tool can work with corporate finance to deal with management fraud detection, credit risk reduction, and corporate finance prediction (Otim, 2012). Data mining is to be considered one of the best tools in addressing the risk of internal fraud in an IT system. Corporate fraud can cause a great amount of money to any organization and it harms the level of security to prevent risk in any IT system as well. By utilizing the data mining techniques in both ways of detection and prevention, fraud risk can be efficiently reduced. In addition, very useful descriptive data mining techniques will help correspond with prediction data mining techniques to effectively reduce frauds and risk in the IT systems of an organization. One of the data mining techniques that involve using prediction and estimation is the motivative latent class clustering algorithm on the descriptive data mining approaches will be very useful in detecting the current risk of internal frauds. According to a case study from Jans's paper, univariate analysis is performed and then multivariate clustering to determine the pattern of procedures in the selected business process. And the patterns apply to most of the observations in the case study and it is important to have domain experts audit observations in minor groups with deviating patterns (Jans, 2010).

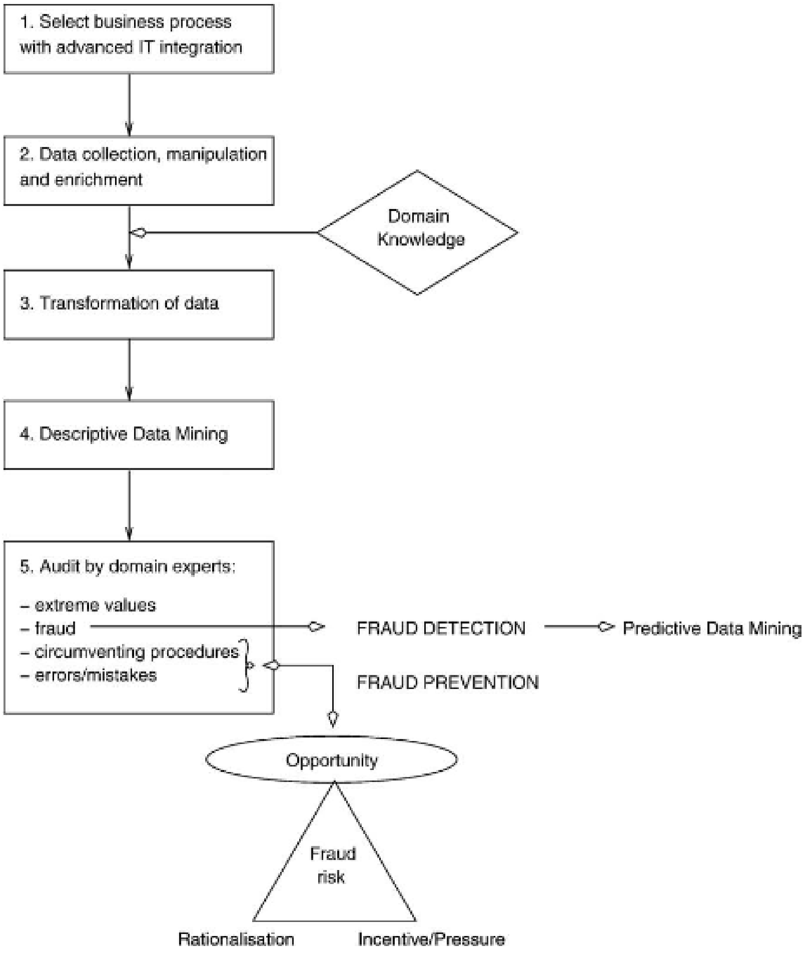

1. Select business process with advanced IT integration
2. Data collection, manipulation and enrichment
Domain Knowledge
3. Transformation of data
4. Descriptive Data Mining
5. Audit by domain experts:
– extreme values
– fraud
– circumventing procedures
– errors/mistakes
FRAUD DETECTION
Predictive Data Mining
FRAUD PREVENTION
Opportunity
Fraud risk
Rationalisation
Incentive/Pressure

Fig. 6. The IFR² Framework as a methodology for Internal Fraud Risk Reduction, from Jan's paper on International Journal of Accounting Information Systems 11 (2010) 17–41.

As we can see from the framework above, the first step is for an organization to select a business process that is important to the organization's development and should be worthwhile investigating. For instance, if the organization has an unstructured business process that involves a large amount of cash flow, or the business process remains unclear and is not well understood by the organization, it will be a good idea for the firm to start an investigation in such business processes. The following step is to gather, manipulate, and enrich the data. This process includes organizing the data in the desired structure and format and creating extra attributes for the data analysis by combining and transforming the original attributes. This step heavily relies on the technical transaction performance by the data analyst of the organization. Then experts with domain knowledge will help with the data translation process. It requires the technical data in the process to be translated into behavioral data. Moving forward, univariate analysis and the multivariate clustering method will be performed to gain more insight from the behavioral data. The domain experts will then categorize the potential frauds from the four groups, which are extreme values, fraudulent cases, cases of circumventing procedures, and errors or mistakes. Fraud prevention in this methodology is primarily based on assessing and minimizing fraud opportunity which is the only element of fraud risk that an employer can influence (Jans, 2010). The results indicate that the use of the framework and the multivariate latent class clustering technique with the data mining approach can be one of the powerful tools to reduce fraud risk in internal frauds and to protect the data safety of the systems. Furthermore, data mining techniques can also be elaborated as a decision support tool to build a better audit support system in organizations. However, there are still concerns about the framework's applicability and accessibility to other different projects in real-world solutions. More research or investigations should be implemented in this area to further prove the reliability and accuracy of using this approach.

7. Monte Carlo Simulation

Many useful and powerful computational simulation methods in the IT system to reduce or minimize risk are introduced in the last section, this section will focus on discussing one of the widely used simulation methods, the Monte Carlo Simulation. As we already know, enterprises or organizations have paid a great amount of money and resources in developing their information systems, due to the rapid growth of technology in the digital era. According to Yang's paper, information system projects nowadays still have a very high rate of failure because of improper risk management. The risk category and risk assessment are widely discussed, however, the use of modeling and simulation on developing information system risk management has been paid little attention (Yang, 2012). Thus, it is important to understand the benefit of using simulation models to develop risk management and assessment plans so that IT system risks such as failure rate can be reduced. The use of Monte Carlo Simulation is very popular in evaluating all kinds of uncertainties and risks and it has become an important tool for risk assessors. The Monte Carlo Simulation is also used for random simulation so it is also named Random sampling technology. It can be applied in solving many different problems in various fields including mathematics, physics, engineering technology, and production management. A probability model or random process with parameter equals its solution should be established first to develop a Monte Carlo Simulation model. Then, a sampling test will be used to calculate the statistical characteristics of parameters based on the model or process. And the last step will be to provide the approximation of the problem and present the solution accuracy with statistical results in the form of the standard error of the approximation. Because of the rapid development of computer technology in the digital era, this method has been widely applied in more areas than ever. Yang’s paper introduced his approach in using Monte Carlo Simulation to identify key risk factors in the development process of an IT system and to evaluate the performance of the system. Yang also identifies the probability distribution of each risk factor and creates formulas and equations between the project performance and the risk factors. Then the Monte Carlo Simulation will be applied to simulate project performance and to produce related frequency charts and sensitive graphs. A software called the Crystal ball is introduced and it is used to fit probability distributions of each risk factor. Crystal ball software is a tool that is good for forecasting, optimization, and risk analysis. It is built in the working environment of Microsoft excel. The first step of using the software to conduct stochastic models is to build up a spreadsheet with edited input variables, output variables, values, and defined models. The next step will be to set up input variables to the assumption and to define the distribution function to each input variable. The output variables will then be sent to the forecasting cell and the mathematical formula will be defined for the modeling. The number of trials will also need to be defined since it indicates the number of randomly selected simulations that will be conducted as the final result. It is also an option to select the sensitive analysis function in the application so that sensitive charts will be generated. And the final step will be to run the simulation and analyze the predicted results from the generated statistical outputs and graphs.

One of the key steps in the process of Monte Carlo Simulation is to develop mathematical models or formulas between the dependent variables and the independent variables. In the case study of Yang’s paper, Structural Equation Models (SEM) are created to establish the regression equation between the project performance and defined risk variables. SEM is a multivariate technique that combines the attributes of both factors’ analysis and multiple regression to simultaneously estimate a series of dependent relationships (Yang, 2012). It is a very powerful model since it can accommodate correlated independent variables and correlated errors. Before choosing the probability distribution of the input variables, it is critical to understand the characteristics and the attributes of the risk factors as well as the empirical dataset. And by comparing the parameter estimates and the significance figures, the goodness of the fitted model can be interpreted. If the P-value of the risk factor is greater than 0.05, this indicates that the zero-hypothesis is failed to reject and the hypothesized model is considered to have a good fit with the sample data. In summary, the SEM method is useful to establish multiple regression equations between the project performance and the determined risk factors. Before performing the Monte Carlo Simulation, the data set will be fitted with a probability distribution for each risk factor and frequency charts will be generated for analysis purposes. The mean value, standardized errors, and significant figures will be measured to consider whether the model indicates a good fit or not. Sensitive charts are useful to determine which risk factors in the model have the greatest positive or negative impact on the project performance.

8. Future Work and Conclusion

Many great approaches of using big data, data analysis, and data processing to reduce IT system risks were introduced in this paper. Due to the length of this paper, not all methods are discussed. As we all know, digital technologies are growing rapidly and driving transformative changes in every aspect of the industry. Organizations or firms take advantage of their transformative changes by refining the way to create, deliver, and capture the value from these changes. Meanwhile, the organizations need to identify and understand well the new risks that are associated with the new transformative changes. Proper actions or effective risk management plans should be implemented to address those identified risks so that more business value can be derived from this effect and the IT system of the organization can be better protected. What's more, a more balanced view of digital technologies between digital transformation and risk management can be performed because both the source of the risk and the management of the risk can be controlled. However, some of the organizations may encounter new risks that they have not recognized before when emerging new technologies in their fields of work. And sometimes these risks can be difficult to handle and may add more complexity to the existing risk. Therefore, tackling these risks concurrently for the organization is the key since most of these risks may have some of their characteristics in common. Organizations and their risk management teams should be adept with these new risks quickly by creating new approaches to how they view risk, manage risks, and harness risk of growth. And by doing this, the value of their digital investments can be captured and the newly identified risks can also be addressed.

It is always an argumentative topic on the thoroughness of the risk management plans whether both quantitative and qualitative approaches should be incorporated. Some people would argue to only use the most feasible approach as the method to apply with risk management. According to Shahidi’s paper, quantitative analysis can be an expensive and/or time-consuming approach that is typically reserved for only a small subset of projects (Shahidi, 2011). Also, discrete problems are often attached with those risks in the system which the inter-dependencies effects are very likely to be ignored. Therefore, there is a need to first understand the nature of the risks and apply the appropriate risk management plans to address those issues. System dynamics is a very powerful method that can be applied in such a situation where it can help people better understand the complex interaction within a system. Throughout the past couple of decades, the use of system dynamics has successfully addressed problems such as understanding the spread of disease, modeling power-grid systems, and dealing with delay-and-disruption claims. System dynamics is also applicable for large and complex projects. There is also a part of the method called the System Thinking approach which focuses on helping to understand the complex interactions in the project risk. This method pays more attention to the uncertainties of the projects rather than the system itself. It can model all the dynamics of the system under scrutiny incorporated with the concept in both quantitative and qualitative analysis into a single model. For instance, the use of system dynamics can have an impact on business analysis since the data used in the analysis process can be simulated and a variety of the variables can be manipulated and controlled to support the strategic decision-making process. System Thinking has a great impact on describing the performance of the system by either reinforcing or reducing several risk variables in the system. This involves the reduction of a set of flows in the system which are interactive with each other. However, there may be risks that can be added to the model which impacts the overall costs and drive the changes in the flows. To overcome this issue, the conducted model in the system dynamic method will be simulated over the apparent duration of the project and the statistical results will be compared with expectations. In addition, the model will also examine the identified risks which have a disproportionate impact on the goals or the objectives of the project. A catastrophic failure will also be detected even if the risk factors are combined with other risks in the system. However, there are still some limitations in system dynamics since the technical requirement is relatively high. Project managers are recommended to obtain training in the field of System Thinking and Dynamics as a way to improve the rate of success in projects. The paper also suggested more research should be performed to further prove that this method can become the best alternative tool for all projects in reducing risks.

Many risk reduction methods are well discussed in this paper and most of them involve using advanced technology and computational simulations. As is mentioned earlier that many new risks can be produced during the process of risk management and assessment. The concern on data usage, transparency, security, and accuracy is raising on the development of the IT system in every single industry. The ability for the organization to obtain a balance between its data security and its right to data remains hard to achieve. The increased number of users expected to have control over their data including accessibility and the right to limit their data makes it more difficult for the firms or organizations to ensure their IT system's safety. Furthermore, if the organization promotes tough strategic decisions which may potentially back away from the data monetization and may have negative impacts on the trust of the customers and stakeholders. This is an important aspect that organizations should be careful with since the relationship between customers and stakeholders is very important. One action that the organizations should take is to implement data assessment meetings and plans to align the proposed data usage and other data information with the organizational value, the expectation from the stakeholders, and the regulatory restrictions. It is also critical for organizations to always think about managing their brand and credibility in the market by focusing on preserving and protecting the value of their critical and sensitive data. Through applying advanced technical methods, important data can be safeguarded since these methods involve AI-enabled controls and many other advanced surveillance methods to mitigate risks. Even though it is hard or impossible to eliminate all the risks in a system, understanding and treating data as a valuable asset and identifying all the associated risks and opportunities can be a good start for an organization to reduce risks or threats.

9. References

Cheraghi, E., Khalilzadeh, M., Shojaei, S., & Zohrehvandi, S. (2017, December 14). *A mathematical Model to select the Risk Response Strategies of the Construction Projects: Case Study of Saba Tower*. Procedia Computer Science. https://www.sciencedirect.com/science/article/pii/S1877050917322809.

Cole, D., Nelson, J., & McDaniel, B. (2015). Benefits and Risks of Big Data. https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1037&context=sais2015.

Jans, M., Lybaert, N., & Vanhoof, K. (1970, January 1). *Internal fraud risk reduction: Results of a data mining case study, by*. International Journal of Accounting Information Systems. https://ideas.repec.org/a/eee/ijoais/v11y2010i1p17-41.html.

Jasiul, B., Szpyrka, M., & Śliwa, J. (2014, December 19). *Detection and Modeling of Cyber Attacks with Petri Nets*. MDPI. https://www.mdpi.com/1099-4300/16/12/6602.

Koen, H., & Koen, H. (2016, May 2). Risk Management Monitor. http://www.riskmanagementmonitor.com/how-the-internet-of-things-benefits-risk-management/.

Otim, S., Dow, K., Grover, V., & Wong, J. (2012, July). *The impact of information technology investments on downside risk of the firm: alternative measurement of the business value of IT*. The Impact of Information Technology Investments on Downside Risk of the Firm: Alternative Measurement of the Business Value of IT. https://www.researchgate.net/publication/262087633_The_Impact_of_Information_Technology_Investments_on_Downside_Risk_of_the_Firm_Alternative_Measurement_of_the_Business_Value_of_IT.

Shahidi, A., & Gray, M. (2011, October 22). Applying the principles of system dynamics to project risk management or "the domino effect". https://www.pmi.org/learning/library/principles-system-dynamics-risk-management-6186.

Szpyrka, M., & Jasiul, B. (2017, February 26). *Evaluation of Cyber Security and Modelling of Risk Propagation with Petri Nets*. MDPI. https://www.mdpi.com/2073-8994/9/3/32.

Tohidi, H. (2011). The role of risk management in IT systems of organizations. *Procedia Computer Science*, *3*, 881–887. https://doi.org/10.1016/j.procs.2010.12.144

Wu, D. D., & Olson, D. L. (2013, September 6). *Computational simulation and risk analysis: An introduction of state of the art research*. Mathematical and Computer Modelling. https://www.sciencedirect.com/science/article/pii/S0895717713002549.

Yang, W., & Tian, C. (2012, January 16). *Monte-Carlo simulation of information system project performance*. Systems Engineering Procedia. https://www.sciencedirect.com/science/article/pii/S2211381911001925.