Proceeding Forum in Research, Science, and Technology (FIRST) 2016

DATA GOVERNANCE IN THE RENEWABLE ENERGY DEVELOPMENT: ISSUES AND CHALLENGES

Sonny Zulhuda [1)]

[1)]*Ahmad Ibrahim Kulliyyah of Laws, International Islamic University Malaysia, Kuala Lumpur, Malaysia*
E-mail: sonny@iium.edu.my

Abstract. Central to the development of the renewable energy is the optimization of the resources and power despite its intermittent characteristics. To this end, the information and storage system needs to be properly managed and governed. Therefore, ICT and its deployment in the renewable sector becomes a governance issue, not a mere matter of business process or technicalities. This paper looks at the criticality of information system in relation to the renewable energy sector. It also identifies relevant data risks primarily the data security and privacy risks, and then examines the scope required for data governance. This is a work in progress to introduce the issues and challenges in relation to data governance in the renewable energy development.

Keywords**:** Renewable energy; data governance; Big Data; critical information infrastructure; privacy; security

I. INTRODUCTION

One of the universal development agendas outlined by the United Nations General Assembly resolution 70/1, namely, the Sustainable Development Goals (SDGs) is to ensure access to affordable, reliable, sustainable and modern energy for all [1]. This is because energy plays important role in the eradication of poverty through advancements in health, education, water supply and industrialization, to combating climate change [1], [2]. Renewable energy is increasingly seen as the next choice to replace the current energy resources which are primarily based on fossil fuels. Considered efficient to reduce pollution and to create a cheaper resources economy, the US government targeted in 2025 that 25% of the country's electricity shall come from renewable resources such as those derived from hydropower, solid and liquid bio-fuels, wind, the sun, biogas, geothermal and marine sources, and waste [2]. Meanwhile in Germany, the government planned to fully fulfill their power demands using renewable energy sources by 2050 [3]. Globally, it was noted by REN21 [4] that renewable energy provided an estimated 19.2% of global final energy consumption in 2014, and growth in capacity and generation continued in 2015. Taken together, this is a strong sign that a global energy transition is under way towards the renewable energy [4].

However, renewable energy is not without problems and difficulties. The main problems come from its intermittency (as the sun is not always shining, and the wind is noÉt always blowing) and location (which may be in the remote area from where such energy is needed to be) [2]. In order to anticipate this intermittency and location problems, notion to develop effective and efficient power storage facilities and technologies became urgent. This would require a critical infrastructure backed-up by strong data analytics and efficient information systems. While data analytics are necessary to read and understand the pattern of environment so as to ensure the storage facilities work to its optimum, the efficient information system will have to ensure such process runs smoothly without unnecessary interruption. This is where the need for data governance comes into the picture. Efforts must be taken to ensure the operators of the data analytics and information system are backed up with conscious steps to ensure the security, reliability and sustainability of the data system.

This study looks up at why data governance is critical to those who develop renewable energy. It asks three basic fundamental questions, firstly what is that Big Data ecosystem behind the renewable energy which necessitates data governance? Secondly, how is the notion of critical information infrastructure relevant to the data governance? And thirdly, what are the data risks that become the immediate attention for data governance? This paper hypothesizes that data governance is increasingly critical for renewable energy sector due to the optimization of ICT and Big Data in the industry.

II. METHODS

This paper aims to answer the research questions by way of literature study on the latest findings presented in academic journals and publications. This is to be achieved by discussing, firstly, the rise of Big Data surrounding the renewable energy system, secondly, the identification of renewable energy sector as a critical infrastructure and what it means for data governance, and thirdly, the identification of the data security and privacy risks and their governance principles reflected from industrial best practices and legal framework. Parts of Malaysian and UK legal frameworks will be used only as a matter of example, and not indicating the jurisdiction-specific laws. The paper will also conclude by preliminarily suggesting data governance tool and its scope.

A. The rise of Big Data in energy sector

The use of information and communications technology (ICT) in the energy sector steps up tremendously. Big Data and advanced analytics are deemed pivotal in improving our ability to forecast, automate, customise and democratise energy. REN21 noted this growth became more obvious lately with continued advances in renewable energy technologies, ongoing energy efficiency improvements, increased use of smart grid technologies and significant progress in hardware and software to support the integration of renewable energy, as well as progress in energy storage development and commercialisation [4]. With the best data analytics, energy providers can work out how to store and distribute energy generated from solar panels and wind turbines [5].

Grolinger, et. al. (2016) highlighted that the advances in sensor technology and the proliferation of smart metering devices that measure, collect, and communicate energy consumption information have created possibilities for development of sophisticated energy services [6]. Not only that, the Big Data collected by smart energy meters have created opportunities to analyse energy use, identify potential savings, customize heating and cooling activities for savings and comfort, measure energy efficiency investments, provide energy cost estimates for real estate buyers, and educate about responsible energy usage and conservation [6].

Definitions and concepts of Big Data in general have been discussed in many studies [7]–[9]. Big Data refers not only to specific, large datasets, but also to data collections that consolidate many datasets from multiple sources, and even to the techniques used to manage and analyze the data [7]. Big data is big in the quantity and variety of data that are available to be processed to make inferences [7]. Gandomi and Haider (2015) noted that 95% of Big Data are unstructured data [8]. The increasing use of Big Data has generated new challenges mainly on problems such as data processing, data storage, data representation, and how data can be used for pattern mining, analysing user behaviours, and visualizing and tracking data [9]. Arguably, this complexity is also evident on Big Data usage relating to the behaviour patters of energy users.

As noted earlier, renewable energy has some drawbacks including its intermittency. This has lead to high improvements in storage technology [3]. There is a need for the smart technology and smart sensors to collect real-time weather information to help improving the accuracy of energy availability in the near future [3]. This is reflected in the increasing use of smart grid and smart homes for a better renewable energy.

On top of that, Schuelke-Leech, et. al. (2015) stressed that a proper management of those Big Data technologies is crucial [10]. As Big Data provides the opportunity to better monitor, correct, and integrate smart grid technologies and renewable energy, data management and utilization must be integrated into organizational operations if the potentials are to be realized [10].

B. Renewable energy system as a Critical Information Infrastructure

With the escalation of ICT use and Big Data analytics in its development, renewable energy system turns to one of the critical information infrastructure (CII). Critical infrastructure has been defined in various aspects, with most of the definitions referring to nation's public assets. The definition of critical infrastructure terminology has been expanded dependent on its function and size of the critical asset [11]. The 'criticality' refers to the utmost public and national concern crucial to the survivability of a nation. The destruction or disruption of these systems and communication networks would significantly affect the economic strength, image, defense and security, government capabilities to function, and public health and safety [12]. Countries differ on the concern depending on public policy adopted by each country. They may see a criticality when it involves a high-stake national security, economy, public health and safety, essential government services or even a relatively-less-harmful 'national image'. Countries put it in their own perspective depending also on the national needs and circumstances [13].

For examples, the United States defines critical infrastructure as systems and assets, whether physical or virtual, so vital to the United States that the incapacity or destruction of such systems and assets would have a debilitating impact on security, national economic security, national public health or safety, or any combination of those matters [13]. Meanwhile, the UK, puts it as one that may involve a major detrimental impact on the availability or integrity of essential services, leading to severe economic or social consequences or to loss of life. Australia focuses on the social or economic well-being and the security of the nation. On the other hand, Japan terms such criticality to the 'great disruption of people's social lives and economic activities'. Meanwhile, Germany outlines the 'long-lasting supply bottlenecks, significant disturbances in public security or other dramatic consequences,' while Malaysia views criticality on the "severe impact that may be caused to national economic strength, national image, national defence and security, government capabilities to function, public health and safety". Based on the above criteria, it is found that energy including renewable energy falls under a "critical infrastructure" and therefore the information and data systems involved in the renewable energy development are a critical information infrastructure. The summary of these can be drawn in Table 1 [13].

TABLE I
THE SCOPE OF CRITICAL INFRASTRUCTURE

Country	Scope
The United Kingdom (UK)	Communications, emergency services, energy, finance, food, government, health, transportation, water.
The United States of America	Agriculture, food, water, public health, emergency services, government, defence industrial base, information and telecommunications, energy, transportation, banking and finance, chemical industry and hazardous materials, and postal and shipping.
Germany	Transportation and traffic, energy, hazardous materials, telecommunication and IT, finance and insurance, services, public administration and justice system, others (including media,

	major research establishments, cultural assets).
Japan	Telecommunications, finance, civil aviation, railways, electricity, gas, governmental/administrative services (including local governments), medical services, water works, and logistics.
Malaysia	National defence and security, banking/finance, information and communications, energy, transportation, water, health services, government, emergency services, food and agriculture.

What we can note from the above revelation is that the choice of developing renewable energy sector brings not only economic but also social and political impact. Indeed, energy infrastructure is a critical underpinning of modern society that any compromise or sabotage of its secure and reliable operation has an enormous impact on people's daily lives and the national economy [14]. This is reinforced by Scholten and Bosman in their work [15]. They argued that shifting from fossil fuel-based to renewable energy brings about a shift in strategic positioning in infrastructure management and advantage goes to those countries being able to render balancing and storage services [15]. This means that countries who develop renewable energy should consider the criticality of their renewable energy resources and take steps to protect their energy system.

C. Data governance on privacy and security risks

Massive data sets pose challenges due to their size and complexity and managing those challenges is a top priority for data management [6], [10]. Electric utilities are heavily-regulated, with utmost concern on their system reliability and profitability. This leads to the necessity of addressing technological, economic, institutional, and policy constraints [10]. McMillin (2012) argued that this issue of data privacy and confidentiality became a major concern for electric power infrastructure and smart grid systems [16]. More recently, it is pointed out by Zhou, Fu, and Yang (2016) that the challenges of big data-driven smart energy management in IT infrastructure include issues of data collection and governance, data integration and sharing, processing and analysis, security and privacy [17].

Data privacy regulations around the world emphasise a fair information practices, where, among others, personal information must not be processed, disclosed or shared unless with consent from the data subject. Besides, those personal information must be managed securely, not being subject to unauthorised access or threats to its integrity [18]. Data security, on the other hand, requires the preservation of the confidentiality, integrity and availability of data [19].

TABLE III
MALAYSIAN PERSONAL DATA PROTECTION PRINCIPLES

Principles	**Scope**
General principle	No process of personal data which is excessive and/or without the consent of data subject.
Notice and Choice	Proper notification must be served on the purpose, etc of that data processing
Disclosure	Unauthorized disclosure or sharing of personal data is prohibited
Security	Technical and organisational measures are imposed to secure personal data
Retention	Personal data shall not be kept longer than necessary
Data Integrity	Data users must ensure that personal data is correct, complete and updated
Data Access	Data users must allow data subject to have an access to his own personal data

In a smart grid, networking, intelligent communications technology and information processing functions are immersed into every facet of the energy system ranging from power generation, power transmission, power distribution and consumer appliances. Hu and Vasilakos (2016) identified new challenges in data analytics in relation to the use of smart grid. They noted that smart energy data contain individual user's private information which is required to be protected under various legal regulations [20]. The data can also contain sensitive information of an organization, which can be used to make decisions affecting the safe operation of the critical infrastructure. Therefore security and privacy will be an important issue. However, this is also very challenging due to the big data nature of the smart energy data, tight cyber-physical couplings, distributed and open environment of the infrastructure [20].

Meanwhile, data security is defined and characterized by several items, including data reliability, confidentiality, availability, integrity, maintainability, authenticity and non-repudiation [20], as showed in Table III:

TABLE IIIII
THE PROPERTIES OF INFORMATION SECURITY

Property	**Scope**
Reliability	Even under the disturbances of faults and cyber-attacks, the system correct service can be maintained within a certain level.
Confidentiality	Data or information is not made available to unauthorized persons or processes. In the proposed dependability framework, it refers to the property that unauthorized persons or processes will not be able to observe the values/contents of the sensitive variables of the relevant systems.
Availability	Readiness for correct service. The correct service is defined as delivered system behavior that is within the error tolerance boundary
Integrity	Absence of malicious external disturbance, which makes the system output off its desired service
Maintainability	Ability to undergo modifications and repairs
Authenticity	Ability to provide services with provable origin
Non-repudiation	Services provided cannot be disclaimed later

Given the strategic value of the CII to the greater interests of the country's economy and security, the CII has been subject to diverse threats and breaches. One of the big threat is data system leak which may lead to system interruption, data theft as well as data destruction [19], [21]. Another attack may be to disrupt partially or entirely the system

attached with the CII as part of cyber-terrorist agenda [20]-[22]. Other threats may emerge as a result of natural disaster [23].

In relation to the energy system as a CII, Arghandeh et.al. (2016) found that, due to the modern society's heavy reliance upon a complex energy smart system which includes advanced sensors, intelligent automation, communication networks, and information technologies (IT), more interconnections and interdependencies between the physical and cyber components of the grid emerge [23]. Therefore, the vulnerability of infrastructure networks to electric grid outages is becoming a major global concern. In order to minimize the economic, social, and political impacts of large-scale power system outages, the grid must be resilient in addition of being robust and reliable [23]. Besides, it has been obvious that cyber-terrorists could have exploited the vulnerability of the energy system [24], [25].

The major data privacy and security concern must be managed and handled properly by those involved in the development and provision of renewable energy to basically maintain the critical functionality of its system backbone [23]. Such concern must not only be felt by people on the operational level, but must also be considered by greater public [26]. This leads to the need for a sound and proper data management and data governance [17]. Nevertheless, this often has not been considered as carefully as they ought to be [25]. Kloza et.al. (2015) argued that the smart energy system must determine a governance that promises an adequate response to environmental, societal and technical challenges [27]. Such governance necessarily considers the regulatory framework surrounding the issue of data management, namely data protection law. Hence, Kloza et. al. Further argued that one has to consider including a data protection impact assessment framework as it can prove to be an effective impact assessment of emerging technologies satisfying certain quality criteria [27]. This has also been supported by Seto (2015) who discussed the importance of privacy impact assessment (PIA) concerning the privacy risks associated with advances in the standardization of the smart grid [28].

PIA is more than just audit and compliance checklist. It covers a whole data life-cycle process, from planning, strategizing to execution [29]. It is a part of a system of incentives, sanctions and review, and should be embedded in project workflows or quality assurance processes. It aims to expose and mitigate privacy risks, to avoid adverse publicity, to save money, to develop an organisational culture sensitive to privacy, and also to build trust and assist with legal compliance [29]. Based on this, PIA is an alternative tool for renewable energy developers to consider in governing and managing their data assets.

III. Results

Based on the study above, the findings can be summarised as follows: Firstly, it is found that the Big Data ecosystem which increasingly supports the renewable energy relies very much on the data analytics. This data analytics can only be successful if the Big Data is properly managed and well governed. Secondly, the adoption of renewably energy sector fits the nature of critical infrastructure due to its critical value for the national economy and security. Therefore, data systems behind the renewable sector will consequently have to be afforded a critical information infrastructure protection. Thirdly, the study discovers that there are many challenges surrounding the sustainability of the data system behind renewable energy, and among those major ones are the protection of data privacy and data security.

Therefore, this paper finds it true that data governance is critical for renewable energy development due to the optimization of ICT and Big Data in the industry, and such governance should consider all the risks assessment and measurement to be taken by all relevant stakeholders, taking into account the necessary technology (tools) and processes (law, best practices, etc.). As such, this paper draws the scope required to make data governance effective as in the Fig.1 below.

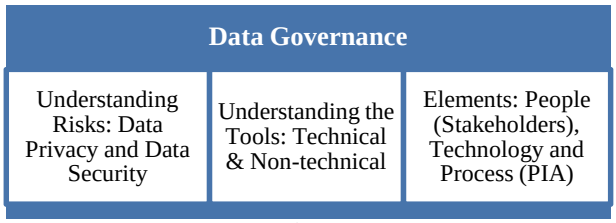

Fig. 1 Scope of Data Governance

IV. Discussion

A. *Recapitulation of Findings*

The preceding paragraphs of this paper gives us an insight that renewable energy development does not only involve questions relating to the turbines and fuels, but also about how well the system and data are being managed and protected. The intermittent nature of renewable energy makes it mandatory to have such a reliable and secure data system that would optimise the storage, distribution and usage of the energy. The role of ICT is no longer restricted to being equipment and tools, but defines the level of sustainability and governance.

Moreover, renewable energy has becomes the next big thing as it is expected to substantially replace the consumption of traditional fossil fuel-based energy. Therefore all industry and infrastructure behind it would eventually become a critical infrastructure, capable of determining a considerable level of national security and economy in near future.

In this respect, risks assessment and risks management are inevitable. This paper highlights that threats to data privacy and data security are among the major challenges to address.

B. *Implications of Findings*

The study discloses few steps to be done by everyone involved in the development of renewable energy system including policy-makers and the market players.

When governments outline their national policy or blueprint on renewable energy, they must not leave out the necessity to also develop and enforce data governance framework. The mandate should also include the requirement of data risks assessment and risk management to be taken by the operators of renewable energy systems. As

for the industry players, they must ensure that any "smart" initiatives to be taken in near future should consider the risks to their data system. Implications to data risks such as privacy and security shall not be an afterthought. They must be inherently embedded, or be designed by default.

The study also reveals that the data governance principles must consider the existing legal and regulatory framework of a country together with any best practices available in the industry. Increasing regulations on data privacy and information security such as laws on personal data protection, confidential information, patent and intellectual property rights, computer crimes law as well as electronic commerce laws would be necessary in understanding the set of data governance required to be complied with.

C. Future Research

The paper nevertheless has limited itself to only issues of data privacy and data security, while there are other aspects of data governance that also require attention and protection, which would require subsequent research in future. Moreover, further research needs to be done to investigate any special nature of data governance when it comes to the renewable energy sector.

V. CONCLUSIONS

Renewable energy promises many things for the sustainable development goals. Such promises can only be realized if the development is made as a comprehensive agenda, taking into account all aspects including the data governance issues and challenges. Nevertheless, data governance is not a one-off product, but rather, a continuous journey.

REFERENCES

[1] United Nations Economic and Social Council, "Progress towards the Sustainable Development Goals," *Report of the Secretary-General No. E/2016/75*, 5th July 2016.

[2] G. Heal, *The Economics of Renewable Energy*, NBER Working Paper No. 15081, 2009.

[3] M. Jaradat, M. Jarrah, A. Bousselham, Y. Jararweh, and M. Al-Ayyoub, "The Internet of energy: Smart sensor networks and big data management for smart grid," *Procedia Computer Science*, vol. 56 (2015), pp. 592 – 597.

[4] Renewable Energy Policy Network for the 21st Century (Ren21), *Renewables Global Status Report*, 2016.

[5] H. Sheffield, "Why energy may soon be free thanks to solar, wind, storage and big data," *The Independent*, 4 August 2016.

[6] K. Grolinger, A. L'Heureux, M. A. M. Capretz, and L. Seewald, "Energy forecasting for event venues: Big data and prediction accuracy," *Energy and Buildings*, vol. 112 (2016), pp. 222–233.

[7] S. Zulhuda, I. M. A. G. Azmi, and N. Hakiem, "Big data, cloud and "bring your own device": How the data protection law addresses the impact of 'datafication'," *Advanced Science Letters*, vol. 21 (10), 2015, pp. 3347-3351.

[8] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics." *International Journal of Information Management*, vol. 35, Issue 2, April 2015, pp. 137–144.

[9] G. Bello-Orgaz, J. J. Jung, and D. Camacho, "Social big data: Recent achievements and new challenges," *Information Fusion*, vol. 28, March 2016, pp. 45–59.

[10] B. Schuelke-Leech, B. Barry, M. Muratori, and B. J. Yurkovich, "Big data issues and opportunities for electric utilities," *Renewable and Sustainable Energy Reviews*, vol. 52, December 2015, pp. 937–947.

[11] E. Hamzah, M. N. B. Jaafar, and H. B. Ali, "Identifying the criteria for critical infrastructure selection," *Proceedings of the 26th International Business Information Management Association Conference - Innovation Management and Sustainable Economic Competitive Advantage: From Regional Development to Global Growth, IBIMA 2015*, 2015, pp. 3400-3406.

[12] Z. Yunos, R. Ahmad, S. H. Suid, and Z. Ismail, "Safeguarding Malaysia's Critical National Information Infrastructure (CNII) against cyber terrorism: Towards development of a policy framework," *6th International Conference on Information Assurance and Security, IAS 2010*, pp. 21-27.

[13] S. Zulhuda, "National security in Malaysia's digital economy: Redefinition, reaction and legal reform," *Journal of Applied Sciences Research*, vol. 7 (13), 2011, pp. 2316-2325.

[14] H. Qi, X. Wang, L. M. Tolbert, F. Li, F. Z. Peng, P. Ning, and M. Amin, M., "A resilient real-time system design for a secure and reconfigurable power grid," *IEEE Transactions on Smart Grid*, vol. 2, Issue 4, December 2011, pp. 770-781.

[15] D. Scholten and R. Bosman, "The geopolitics of renewable: Exploring the political implications of renewable energy systems," *Technological Forecasting and Social Change*, vol. 103, February 01, 2016, pp. 273-283.

[16] B. McMillin, "Privacy and confidentiality in cyber-physical power systems," *IEEE Power and Energy Society General Meeting*, 2012, Article number 6345667.

[17] K. Zhou, C. Fu, and S. Yang, "Big data driven smart energy management: From big data to big insights," *Renewable and Sustainable Energy Reviews*, vol. 56, April 2016, pp. 215–225.

[18] A. B. Munir, S. H. M. Yasin, and M. E. Karim, *Data protection law in Asia*, Sweet & Maxwell, Hong Kong, 2014.

[19] S. Zulhuda, "The state of e-government security in Malaysia : reassessing the legal and regulatory framework on the threat of information theft." In: *1st Taibah University International Conference on Computing and Information Technology (ICCIT 2012)*, 12-14 March 2012, Madinah, Saudi Arabia.

[20] J. Hu and A. V. Vasilakos, "Energy big data analytics and security: Challenges and opportunities," *IEEE Transactions on Smart Grid*, vol. 7, Issue 5, September 2016, pp. 2432-2436.

[21] I. M. A. G. Azmi, S. Zulhuda, and S. P. W. Jarot, "Data leak, critical information infrastructure and the legal options: What does Wikileaks teach us?" *International Journal of Cyber-Security and Digital Forensics*, vol. 1 (3), 2012, pp. 226-231.

[22] K. S. Stegen, P. Gilmartin, and J. Carlucci, "Terrorists versus the sun: Desertec in North Africa as a case study for assessing risks to energy infrastructure," *Risk Management*, vol. 14, issue 1, February 2012, pp. 3-26.

[23] R. Arghandeh, A. Von Meier, L. Mehrmanesh, and L. Mili, "On the definition of cyber-physical resilience in power systems," *Renewable and Sustainable Energy Reviews*, vol. 58, May 2016, pp. 1060-1069.

[24] J. Lilliestam, "Vulnerability to terrorist attacks in European electricity decarbonisation scenarios: Comparing renewable electricity imports to gas imports," *Energy Policy*, vol. 66, March 2014, pp. 234-248.

[25] J. Kent, J. Rowe, and C. Tomlinson, "Briefing: Information security in community energy," *Proceedings of Institution of Civil Engineers: Energy*, vol. 169, Issue 2, 1 May 2016, pp. 48-51.

[26] T. Van de Graaf and B. K. Sovacool, "Thinking big: Politics, progress, and security in the management of Asian and European energy megaprojects," *Energy Policy*, vol. 74, Issue C, 2014, pp. 16-27.

[27] D. Kloza, N. van Dijk, and P. De Hert, "Assessing the European approach to privacy and data protection in smart grids. Lessons for emerging technologies," *Smart Grid Security: Innovative Solutions for a Modernized Grid*, August 12, 2015, pp. 11-47.

[28] Y. Seto, "Application of privacy impact assessment in the smart city," *Electronics and Communications in Japan*, vol. 98, Issue 2, 1 February 2015, pp. 1427-1435.

[29] Information and Communications Office (UK). *Privacy Impact Assessment Handbook*, ICO, Londong (undated).