Leveraging dark data for governance of Kenya Electricity Transmission Company

Gloria Regina Mukhongo
Moi University, Kenya
gloriamukhongo@gmail.com
https://orcid.org/0000-0001-9037-2948

Evans Munge Mwangi
Moi University, Kenya
munge.mwangi.schola2@gmail.com
https://orcid.org/0000-0001-8500-282X

Elsebah Maseh
Moi University, Kenya
emaseh@mu.ac.ke
https://orcid.org/0000-0001-6975-663X

Received: 06 March 2024
Revised: 03 June 2024
Accepted: 18 August 2024

Abstract

Dark data entails information assets organisations collect, process and store during regular business activities, but fail to use for any other purposes. Yet it has the potential to be a very powerful tool for informing business decisions. Dark data continues to be a risk at the Kenya Electricity Transmission Company Limited (KETRACO). This is seen as an embarrassing exposure in the media when it lands in foreign hands such as bloggers, lawyers, and auditors, but can hardly be traced by the company itself. This has been blamed on the lack of proper structures for data management. Therefore, this study sought to establish how dark data can be leveraged for effective governance in KETRACO. The objectives of the study were to present the context of dark data creation and its capturing at KETRACO; and to establish how dark data can be leveraged for effective governance in KETRACO. The study adopted a case study research design within a qualitative approach. For that purpose, personal interviews were used as the primary tool for data collection from 32 participants. The findings of the study revealed that dark data at KETRACO accumulated in different locations and formats with no specific individuals responsible, as it streamed into the company from both external and internal sources. The findings also showed that dark data comes with risks and opportunities that have implications for corporate governance. Therefore, the study recommended that organisations such as KETRACO should implement clear strategies to maximise the opportunities presented by dark data. This study is important to the field of information science because of its emphasis on leveraging all organisational data, and by extension records for effective governance.

Keywords: dark data, information, data management, governance, corporate governance, effective governance, Kenya Electricity Transmission Company

https://dx.doi.org/10.4314/jsasa.v57i1.5
ISSN: 1012-2796

1. Introduction and background

Organisations generate or receive data in their daily transactions, which makes it an essential element for decision-making. Data is described by Simplilearn (2023) as texts or numbers written on papers, bytes and bits inside the memory of electronic devices, or facts that are stored inside a person's mind for reference or analysis. As organisations continue to swim in this ocean of varied types of data; current, ancient, structured, unstructured, sensitive and trivial, it is necessary to know whether all this data is being used, and in what manner. It is also key to determine the data that can translate to records for documentary purposes. However, not all this data is used beyond the reason they were created or acquired for. This gives rise to dark data. Gartner Inc. (2023) and KDNuggets (2022) describe dark data as all the information that a business generates, collects, stores and never uses again. Neha and Pahwa (2020:467-474) accentuate dark data as the data that an organisation collects but fails to harness which then forms the biggest portion of big data.

Also referred to as dormant and inactive, dark data crops up when data is collected, analysed and kept as part of normal business operations and protocols, but not used for other functions such direct monetising (Cubeware GmbH 2021). From the foregoing, dark data refers to data that organisations have in their possession, without the ability to access it or with limited awareness of what it represents despite the data's probability of having business-critical information (Cafarella et al. 2016). Splunk (2019) emphasises that dark data is all the unused, unknown and untapped data across an organisation. Data is generated as a result of users' daily interactions online with countless devices and systems, from machine data, server log files to unstructured data derived from social media. However, Crown Records Management (2019) notes that not all dark data is digital; people leave paper everywhere and the ultimate dark data may be stored in a third-party warehouse that businesses are billed for monthly but have almost been forgotten. Banafa (2015) identifies the following three types of dark data:

i. Data that is not currently being collected, therefore one cannot take advantage of it.
ii. Data that is being collected but is difficult to access at the right time and place.
iii. Data that is collected and available, but that has not yet been productised.

DFIN (2022) describes dark data sources as follows: Unstructured sources of dark data entail unprocessed data stored in their native format such as email correspondences. Semi-structured data sources comprise data with some metadata, thus making it searchable or catalogable; and structured sources that contain formatted data such as customer support information. Steel (2022) categorises dark data into three: Critical business data containing highly valuable information relevant to a business's continuous growth and the meeting of goals; redundant, obsolete, and trivial (ROT) data hides in internal networks which once discovered, can be marked for deletion or moved into remediation workflows; and dark data companies never find out they have, yet it poses a constant risk.

Dark data is increasingly recognised globally, but challenges persist (Singh & Miller 2024), particularly in developing countries such as Kenya. Overall, dark data awareness is low in East Africa due to its novelty and the fragmented record-keeping legacy of colonialism. Resource constraints further hinder management compared to developed countries with better infrastructure, established practices and trained professionals. Most developing countries lack expertise and technology to effectively handle this data type (Rao, 2018). Despite these

challenges, international collaboration and local initiatives indicate a growing interest in exploring dark data's potential, despite limited resources and prioritising basic archival functions (atlan.com, 2023).

2. Dark data and governance

Governance has been defined as structures and processes that are designed to ensure accountability, transparency, responsiveness, rule of law, stability, equity and inclusiveness, empowerment and participation (UNESCO-IBE 2023). It entails the norms, values and rules for managing public affairs in a transparent, participatory and receptive manner. Good governance is an approach to governance committed to creating a system founded in justice and peace that protects individuals' human rights and civil liberties. Adding his voice to good governance discourse, United Nations (2024) and Donohue (2022) explain that for an organisation to claim it is practising good governance, it should conduct public affairs and manage public resources transparently while following the rule of law. This results in minimised potential for corruption; increased inclusion and ability to benefit from diverse thinking; and response to the needs of society, both now and in the future (Donohue 2022).

Data is critical for decision-making. York (2023) concurs that there is near-universal understanding that data drives everything from product development and supply chain to customer experience and overall business strategy. This data may include past employee records, transaction logs, confidential surveys, emails, downloaded attachments and even surveillance video footage. However, Raghavan (2023) observes that while organisations have become adept at collecting large volumes of data, they still struggle with leveraging all that data, making 50% of it considered dark. With thorough analysis of dark data, organisations can gain new insights into perpetual challenges and needs of stakeholders, identify new opportunities and make data-driven decisions that improve efficiency and drive outcomes and goals (Hanvey 2023). Bigelow (2021) adds that this data that businesses collect and store but overlook, ignore, forget or simply underuse is called dark data. This data could rather be leveraged in time for effective governance. However, its potential is mostly discovered when it falls into foreign hands such as hackers and auditors, or wrong hands who are prone to interpret it in a negative, skewed way, regardless of whether it is incorrect or incomplete.

Dark data may therefore be one of an organisation's biggest untapped resources and competitive organisations need to tap into its full value (York 2023). Van Loon (2023) and KDNuggets (2022) assert that dark data might be personal, unstructured, regulated or unguarded, thus being an unprotected risk to security and compliance. Hoarded dark data, whether paper-based or electronic, can lead to conclusions and actions that are mistaken, dangerous or even disastrous (Hand 2020). From the foregoing, dark data affects governance in the following ways:

2.1 Breach of confidential and sensitive information

Dark data contains information related to a particular company. Therefore, Soni (2022) explains that if it lands in the wrong hands, it can cause serious damage. Sonawe (2023) and ManageEngine (2023) agree that dark data might store personally identifiable information, intellectual property and business strategies. When both location and contents of files are

unknown by owners, the data is inadequately protected, making it more prone to hackers (Mann 2022). This could result in significant business disruptions, legal action, pay-outs and even lost businesses (KDNuggets 2022; Sonawe 2023).

2.2 Regulatory compliance

Dark data may contain information subject to legal and regulatory requirements, such as privacy laws and industry-specific regulations (Sonawe 2023). However, due to its nature, many organisations may be unable to provide the appropriate security because they are unaware of its content or location. This easily enables theft and leakage of sensitive data or delays retrieval of needed information. This does not only damage an organisation's reputation but also leads to non-compliance with the regulatory requirements, a factor that may attract penalties such as having to pay financial compensation to the affected parties, which causes the company financial losses (Sonawe 2023, KDNuggets 2022; DFIN 2022).

2.3 Storage space and cost

Dark data is inactive information held in the organisation's data ecosystem because of collecting, processing and storing data for single-use purposes never to be reused, such as multiple near-identical images in Google Photos and data from Internet of Things (IoT), among others (World Economic Forum 2023). Steel (2022) highlights that businesses spend millions collecting or analysing new data to derive insights from, while already having relevant information. World Economic Forum (2023) and Srinivasa (2023) observe that this can make organisations incur huge storage and electricity costs and attract hefty non-compliance fees.

2.4 Decision-making

Dark data may contain insightful information that can be used in decision-making. However, Sonawe (2023) explains that it lacks proper governance and quality control measures, resulting in many organisations struggling to retrieve it when needed for decision-making, and if retrieved due to its majorly unstructured nature, it is difficult to trust its accuracy, relevance and reliability. Due to lack of proper data governance, utilising dark data can result in poor decision-making, inaccurate insights and unreliable reporting (Sonawe 2023).

2.5 Missed opportunities

The process of data gathering and storage costs business financially and administratively, which should be compensated by businesses maximising on their data (Sonawe 2023). Labovich (2022) elucidates that dark data can reveal insights into wants and needs of an organisation, including insights around customers and employees, the assets and manufacturing quality and the risks/opportunities its brand faces. However, due to not realising what data is potentially available to them nor the potential of the information in dark data, organisations miss these insights (Yackel 2022). Therefore, Tittel (2014) asserts if organisations decide not to invest in analysis and mining its dark data, concerted efforts by third parties to exploit the value of that data represent potential losses of intelligence and value based on its contents.

2.6 Inefficiency

Dark data can provide insights into internal processes, workflows and employee collaboration (Adusei 2023). However, Sonawe (2023) explains that most of the time, storage of dark data does not involve proper organisation and curation, resulting in concerns such as duplication of data and missing trends or information. These inefficiencies are then seen when organisations need to use decision-making insights from their dark data, but cannot to locate and retrieve specifically needed information, resulting in wastage of time, resources and, ultimately, loss of revenue (Sonawe 2023; Kizen 2023).

2.7 Environmental pollution

The accumulation of dark data is not environmentally friendly. Sarkar (2022) and DFIN (2022) concur that dark data stored away in servers and data banks without much practical use contributes to a large chunk of internet-related emissions. Artikeln, (2023) asserts that dark data is not just a burden on information security and governance, but also responsible for releasing millions of tons of CO_2 into the atmosphere annually. It is key to note that in the East African Community (EAC) region, Kenya and Tanzania, the two largest economies, are also leading in carbon dioxide emission with Kenya's annual emissions in 2021 were 19.88 million metric tonnes, followed by Tanzania with 13.06 million metric tonnes.

2.8 Ease of attack by malware and ransomware

Dark data can be a security risk and a hindrance to operations because organisations are either unaware of its existence or do not know where to find it nor its worth. Therefore, there is little or no understanding of how to protect it from attack nor how to recover it after an attack (Bhatia 2023). Sonaware (2023) asserts that dark data may hold sensitive and valuable information, which may be an enticing target for cybercriminals and ransomware attacks. ManageEgine (2022) underscores unstructured data that is not monitored for sudden spikes in file activities can result in ransomware and other malware infections going undetected.

KDNuggets (2022) acknowledges that dark data is the biggest slice of the big data pie and holds a massive amount of potential for those who can harness it. Van Loon (2023) describes dark data as an untapped treasure trove of potential business value, which if analysed, will unlock important insights as outlined in the following opportunities:

2.9 Customer experience

Some of the sources of dark data include customer feedback and behaviours. These give customer patterns and trends, which if analysed, can give more insight into an organisation's customer (Dash 2022). Labovich (2022) cites customers' and employees' needs, the assets and manufacturing quality and the risks/opportunities your brand faces on social media as some of the insight dark data can give.

2.10 Competitive advantage/Business intelligence

While some organisations consider dark data to be too old to provide value, incomplete, redundant or limited by formats, competitive organisations take it as one of their biggest

untapped resources and assets they need to maximise on (Splunk 2019). Tittel (2014) observes that dark data may encompass proprietary or sensitive information reflecting business operations, practices, competitive advantages, important partnerships and so forth. Therefore, inadvertent disclosure could adversely affect the bottom line or compromise important business activities and relationships.

2.11 Business growth

Dark data represents untapped potential for organisations, which with the right technology tools, can be leveraged to give organisations valuable insights, business opportunities and competitive advantages (Sonawe 2023). As Splunk (2019) asserts, the number of specific dark data use cases are vast, one of the biggest being to create and develop new and more productive enterprise business strategies. Splunk (2019) states dark data can also be used to improve quality assurance; detect and correct errors; and look for privacy loopholes, security vulnerabilities and potential compliance violations.

3. The context of dark data creation/accumulation and storage

Businesses routinely collect vast quantities of data to use for an array of tasks, including business analytics and direct monetisation, yet not all the data collected and stored will be used to benefit the business (Bigelow 2021). Srinivasa (2022) explains that some compliance and governing standards like General Data Protection Regulation (GDPR) force organisations to follow strict regulations for protecting sensitive data. KDNuggets (2022) and Wanyaga, Gikandi and Ndirangi (2017) explain this drive data storage often results in sensitive information being stored in file locations that are unknown long after the mandatory period.

Applications or devices, such as the IoT, collect data by default with little business insight or awareness that the data exists (Bigelow 2021). Cubeware GmbH (2021) illustrates this with online banking where one might think that on their end, they are simply giving login data like username and password, yet they are also generating other data, such as login time and date. Solely stored for regulatory purpose, this data ends up being dark data, as it is not consulted for any other purpose.

Srinivasa (2022) observes that marketing departments, analysts and data scientists collect and store data with an aim of analysing it to inform strategies for effective outreach, such as sales. However, this data is not reused again or by other departments, thus resulting in data silos that are easily forgotten (Srinivasa 2022). Bigelow (2021) state useful data is collected but becomes outdated due to lack of tools and processes to analyse or use all available data. With constant technological advances, companies might lack strategies to migrate their data into new storage software, which in turn accumulates dark data in outdated storages.

Firican (2023) and Authentix (2022) assert that constant advancement in technology with inexpensive storage, facilitates fast accumulation of data. However, the lifespan of this data is short, or they are not cleaned up even after consumption, thus resulting in digital hoarding such as data housed in older, legacy systems. Hand (2020), Corallo et al. (2021) and Rahmouni (2023) agree that over time, new data sources, new ways of collecting data and new types/formats of data emerge, all of which bring with them new types of dark data.

These contribute to the accumulation of dark data due to various formats that are not accessible with the organisation's tools (Van Loon 2023).

Bigelow (2021) notes that dark data, such as contractor details, factory acceptance test reports and feasibility reports may accumulate because of a business projects or initiatives designed to use the data being scrapped, never reaching fruition or losing financial or management support. While awareness of dark data is growing globally, organisations, developing countries like Kenya may lack resources and expertise to manage it. This research can be a valuable case study, paving the way for better understanding of dark data's potential in the African context.

4. Background of Kenya Electricity Transmission Company (KETRACO)

KETRACO is a public company that was incorporated on 2 December 2008 and registered under the Companies Act, Cap 486 of the Laws of Kenya pursuant to Sessional Paper No. 4 of 2004 on Energy (KETRACO 2020). Its core function is to plan, design and construct, own, operate and maintain high-voltage electricity transmission lines and fibre optic cables with the aim of shielding electricity consumers from higher tariffs in future.

KETRACO comprises nine directorates, each with its head, namely: Design and Construction; Project Development Services; System Operations and Power Management; Company Secretary and Legal Services; Human Resources and Administration; Supply Chain Management; Strategy, Research and Compliance; Internal Audit; and Finance. Records Management section headed by a records manager resorts under the Human Resource and Administration Directorate. Records management is decentralised. To execute its mandate, the company acquires and generates a variety of records through the various directorates while carrying out their daily business transactions. Each directorate's secretary manages records with guidance from the records manager (KETRACO 2020). Paper-based records are stored in fireproof safes, bulk filers and roller-shutter cabinets. However, staff also keep some of the paper-based records on their desks, drawers and sometimes at home. The sensitivity of the records influences the storage facility.

From a governance perspective, KETRACO creates and maintains data, records and other information for the following purposes: legal value, for example, easements that serve as agreements between project-affected persons (PAPs) and KETRACO for purposes of subsequent use of affected land and compensation; administrative value, policies giving guidelines on various procedures in the company; fiscal value, like procurement plans that outlines the requirements of the KETRACO and their financial implications to ensure that funds are managed in the most transparent, efficient and cost-effective way; and historical value, example minutes that can be used to explain how the company was formed.

However, not all this data reaches directorates' secretaries, yet it spontaneously accumulates in the dark. This occasions other issues such as necessitating the creation of large storage servers and physical space, exposing data to risks such as unauthorised access that may cause reputational harm to the company while reducing its competitive intelligence. These issues, among others, hinder effective governance.

5. Statement of the problem

KETRACO's inadequate data management processes and unstructured systems resulted in the accumulation of dark data, which has, in turn, been accessed by malicious entities, causing negative consequences and damaging the company's reputation. KETRACO creates and receives lots of data during its daily transactions with stakeholders using different channels. Just as this data is created or received through different channels and by different people, most of it is stored by different people and in different locations. This not only implies that part of KETRACO data is likely to be managed unprofessionally, but also that there is a likelihood of creating a disconnection in data management – a factor that fuels the accumulation of dark data.

The various data sources at KETRACO range from electronic to manual devices. These include personal mobile devices, laptops, closed circuit television (CCTV) cameras, physical files and paper-based documents. This eases data capture and storage at KETRACO, yet there is no elaborate system or mechanisms in place to process all this data for organisational good. These unintended data silos that result from different, unintegrated devices capturing data and the data being managed mostly by the various staff in charge of these devices also result in the accumulation of dark data.

Panorama (2018) likens dark data to a gate crasher at a party, who either turns out to be useful, or costly and risky. This only comes out when it gets into foreign hands that interpret it in a skewered manner mostly to suite them. This has been exemplified in cases of embarrassing exposés of the company in the media as well as on blogs, among other social media platforms. For example, in an article in the *Daily Nation Newspaper,* Kamau (2019) indicated a suspected loss of Sh14.2 billion through alleged fraudulent compensation to landowners by KETRACO. Obura and Wafula (2018) state that KETRACO was fighting a Sh6.3 billion scam. The two cited cases involved suppliers who were allegedly owed millions of Kenyan shillings because of project delays. However, when summoned, for example by the Parliamentary Investigations Committee (PIC), these cases seemed to hit a dead end due to unavailability of required sources of information such as the feasibility report from both parties (Odhiambo 2022). Kenya Insights Team (2022) states that by coincidence or design, the summoned officials were at a loss when they appeared before the committee, and in the end, failed to provide substantive information required, as they all asked for more time to gather all the required documents and prepare coherent statements for presentation. This information was supposedly obtained from dark data sources, including minutes of meetings and closed files.

Such cases not only taint the image of the company but creates stakeholder distrust among other negative repercussions. The two cases exemplify how dark data is a key issue in governance. The ability to track all company dark data and be able to know whenever it is created, who creates it, where and how it is stored and used to inform policy and decision-making is critical. The study, therefore, set out to establish how effective governance in the Kenyan energy sector can be achieved by leveraging dark data. It is believed that KETRACO's inadequate data management processes and unstructured systems resulted in the accumulation of dark data, which has, in turn, been accessed by malicious entities, causing negative consequences and damaging the company's reputation.

6. Objectives of the study

The overall objective of the study was to establish how dark data can be leveraged for effective governance in KETRACO. The study was guided by the following objectives:

i. To establish the context of dark data creation and capture at KETRACO.
ii. To establish how dark data can be leveraged for effective governance in KETRACO.

7. Research methodology

The study was conducted from an interpretive perspective informed by the need to clearly understand the causal relationships between dark data and corporate governance. It adopted a case study research design within a qualitative approach, which allowed the researchers to obtain detailed and meaningful responses. Ninety-one participants were purposively selected from the various directorates based on their information richness from a population of 569. However, saturation was reached upon interviewing participant number 32. The researcher used semi-structured interview schedules as the primary tool for data collection with the aim of obtaining detailed responses. The trustworthiness of the interview schedules was ensured through peer debriefing and member checks. Data obtained from the interviews was analysed using thematic analysis. This involved reading the interview responses/answer sheets to look for repeated ideas, topics or patterns with an aim of finding themes that emerged from the data. The data collected was then presented in narrative descriptions. In order to anonymise the responses from participants, codes were used to represent them. This entailed serialising the interview schedules with numbers from one (1) consecutively, presided with a letter P representing a participant.

8. Results and discussion

The results of the study are provided in the following sections based on the objectives of the study and the themes emerging from data collected.

8.1 Context of dark data

From the findings, the researcher established two main sources of dark data in KETRACO, most notably external and internal sources. The external sources included incoming correspondences such as letters, journals and invoices from professional bodies; Government documents such as Circulars, *The Kenya Gazette*, Constitution of Kenya, visitors (both walk-ins and scheduled ones), external emails, seminars, workshops, webinars and other trainings, internet and incoming calls.

The internal sources comprised data that came from within KETRACO such as staff, work plans, transmission master plan, performance contracts, procurement plans, agreements (contracts, easements, MoUs). policies, standards, manuals, service charter, internal emails, Integrated Land Information System (ILIS), various subject files in the company; calls and peer groups. The findings also indicated that internal dark data was also derived from employment master-files, staff files, various subject files, occurrence books, licences, assets and liability registers and disposal registers. internal data sources also comprised of various

reports, systems applications and products – Enterprise Resource Planning (SAP-ERP), balance sheets, income statements and cash flow statements as their sources of data.

In terms of storage, the study established that dark data was stored manually and electronically. Electronic storage entailed the company's online repositories and shared storage such as SAP-ERP, SharePoint, ILIS, laptops and phones. Most participants cited personal and official mobile devices and laptops as storages of data they required for their daily transactions.

The manual storages included documents/records stored on desks, drawers and file cabinets such as staff notebooks and working files, documents/records in staff homes and tacit knowledge. The findings also showed responsibility of data custody was not clearly defined at KETRACO and there were no defined data storages 'n\or responsible persons.

This can be likened to KDNuggets' (2019) observation that in most organisations, departments have their own data collection and storage processes and systems that may not be known to other departments. The result is that data, even if relevant to other departments, lies unused in the dark in another department.

8.2 Leveraging dark data for effective governance

From the findings, the researcher learnt that dark data had both positive and negative implications for governance at KETRACO. Most participants acknowledged the existence and potential of the dark data, particularly in their possession. Participant P7 said this data was not only consuming the company's and personal electronic and physical space, but financially too. Participant P9 said, *"KETRACO spends a lot of money to store what it is not using."*

The findings showed that dark data at KETRACO held valuable information for the company. Most participants talked about the legacy records referring to this. Legacy records entailed records in outdated formats (this can be obsolete systems, old files, collection of papers or records with no clear organisations) that are difficult to access and are no longer used.

In addition, participant P17 cited phone call transcripts and recordings, saying that "*they could be used to avert some incriminating situations in time or to enhance efficiency*." CIoD, Nigeria (2022) and Corporate Governance Institute (2020) also agree that harnessing dark data not only facilitates objective decision-making but also enables identifying trends, patterns and previously unrecognised opportunities.

The findings depicted dark data as a drive to efficient and responsive governance at KETRACO. Participant P19 agreed that "*the company was prone to miss out on opportunity that could be maximised if dark data is harnessed to assist management to make strategic and efficient decisions based on accurate and actionable insight*". In concurrence with Corporate Governance Institute (2020), analysing dark data will ensure that relevant insights are captured, evaluated and promptly acted upon.

The findings further showed that harnessing dark data enhanced risk management and compliance at KETRACO. Participant P2 said "*Analysed dark data will give a comprehensive view of its data.*" This will enable KETRACO to identify potential compliance breaches, highlight suspicious transactions and mitigate potential risks. Corporate Governance Institute (2020) agrees that harnessing dark data facilitates objective decision-making.

Furthermore, the researcher established that harnessing dark data enhanced transparency and accountability in KETRACO, thus enhancing stakeholder confidence. Participant P2 said, "*By leveraging dark data, there will be more transparency and accountability in Company processes which will enhance stakeholder satisfaction and trust.*" Permutable.ai (2023) agrees that by leveraging advanced analytics, organisations can gain deeper understanding of their stakeholders' needs, preferences and concerns. This knowledge facilitates effective communication, tailored engagement strategies and the alignment of corporate goals with stakeholder expectations. This agrees with CIoD Nigeria (2022) saying with the rise in data-driven decision-making, the importance of transparency and efficiency cannot be overstated, as investors, stakeholders and the public increasingly demanded greater visibility into how companies are managed.

The researcher also found that harnessing dark data would enable KETRACO to anticipate risks and respond proactively. Participant P6 said "Analysis of all data, including dark data, will uncover their risk potential to the company, which will enable the risk to be mitigated." These findings concur with KDNuggets (2022) that the dormant nature of dark data makes it dangerous to the companies/organisations concerned. Van Loon (2023) agrees that dark data is an attractive target for ransomware attacks and data breaches by fraudsters and cybercriminals who leverage this sensitive data for malicious intent, yet only 33% of IT staff and executives are knowledgeable about dark data risks. Hand (2020) adds that hoarding of dark data can lead to conclusions and actions that are mistaken, dangerous or even disastrous.

The findings also depicted dark data as a key player in preserving the rule of law at KETRACO while demonstrating fairness, equity and consistency. Participants P11 and P17 agreed that leveraging dark data ensures transparency and accountability among officers in KETRACO. As a result, this will ensure that company resources are managed appropriately. As Mosweu and Rakemane (2020) demonstrated, by harnessing dark data, there will be adequate information to enable stakeholders to hold public officers accountable for their deeds. As a result, this will ensure various KETRACO stakeholders genuinely get their rights, while reducing the public suffering such as delayed or irregular compensation of PAPs.

Lastly, the findings stated that dark data, specifically manual, is a health hazard. Participants P5 and P7 said the legacy documents storages accumulated dust, which was a potential source of respiratory problems. Participant P16 said that as per the company's Safety, Health and Environmental Policy 2010, the furthest staff can go in this regard is maintaining highest standards of housekeeping and reporting to their immediate supervisor any situation that is likely to present a hazard. Chartered Governance Institute UK & Ireland (2023) and Smith (2022) agree that employee wellbeing entails looking after their psychological and physical wellbeing for purposes of risk management and employee retention because the risks associated with lapses in physical health and safety are obvious, such as lawsuits and hefty personal injury compensations.

9. Conclusion

The research findings indicated that organisations unknowingly invest considerably in dark data while not maximising on its potential. This entails purchasing computers, server spaces and rent for physical data storages, while also paying for their subsequent electricity bills. With the steady advancement in technology, the amount of dark data that organisations create and receive in their daily transactions is likely to keep accumulating in storages, unless something is done. Therefore, this calls for organisations to acknowledge the potential of analysing their dark data to give essential insights and patterns for continuous improvement, compliance and identifying potential risks, among others. To reiterate and as Banafa (2015) emphasises, much dark data remains unilluminated because organisations are unaware of its potential. This means that the Kenyan energy sector and Kenya at large is at a constant risk of losing prospects due to the unused insight hidden in its dark data. This indicates that should proper dark data management be ignored, African countries will continue missing prospective opportunities.

10. Recommendations of the study

The study found that dark data has both negative and positive impacts on governance and therefore recommends that KETRACO management urgently implements policies that focus on the management of dark data both electronically and manually. This will guide staff to be intentional when creating/receiving data and managing data in their possession. The policies should enable staff to distinguish between what data must be collected and stored for compliance reasons, and what does not need to be maintained. This will help KETRACO to create or receive data that is primarily necessary, thus minimising the accumulation of dark data.

While establishing clear dark data governance policies, the study also recommends the need for KETRACO to change how it manages and analyses its data by embracing modern data management techniques such as data analytic tools and employment of data scientists/analysts. These tools will give the company a holistic view of its data from the time of acquisition until the data is disposed of. This will enable the company to understand better its operations, stakeholders and market trends. The study also recommends that KETRACO stakeholders be sensitised continually about dark data. This should entail making them aware of the existence of dark data, how it accumulates, its locations, its impacts and how it can be leveraged. This can be done through conducting webinars or using nuggets on computer screensavers.

Declarations

The authors declare that:

- This manuscript has not been previously published and is not under consideration for publication with any other journal or copyrighted publishing platform of any kind.
- Permission was granted for collection and publication of the presented identifiable data by the National Commission for Science, Technology & Innovation of Kenya (NACOSTI) (Permission letter is provided alongside the manuscript).
- Unlawful statements that infringe any existing copyrights are avoided in the manuscript.

- There is not any potential conflict of interest for the research.
- All authors are familiar with the content of this manuscript and gave consent to co-publish.
- All authors contributed to the writing of the article manuscript.
- Authors take responsibility to keep participants information confidential, as required by legislations including Kenya Data Protection Act, 2019.
- Author(s) gives consent to the Journal of South African Society of Archivist to publish the manuscript.

Acknowledgements

No funding source was involved.

References

Adusei, J. 2023. *The hidden value of dark data*. https://sqagroup.com/the-hidden-value-of-dark-data/ (Accessed 13 September 2023).

Artikeln, A. 2023. *The environmental impact of dark data*. https://www.ironmountain.com/nl/resources/general-articles/t/the-environmental-impact-of-dark-data (Accessed 8 August 2023).

Atlan.com. 2023. *What is dark data and how can you make sense of it?* https://atlan.com/dark-data-101/ (Accessed 17 June 2024).

Authentix. 2022. *Big data analytics (& why dark data matters).* https://authenticx.com/resources/big-data-analytics-why-dark-data-matters/ (Accessed 08 August 2023).

Banafa, A. 2015. *Understanding dark data*. https://www.bbvaopenmind.com/en/technology/digital-world/understanding-dark-data/ (Accessed 08 August 2023).

Bhatia, A. 2023. *Dark data: Managing the data you can't see*. https://venturebeat.com/datadecisionmakers/dark-data-managing-the-data-you-cant-see/ (Accessed 14 September 2023).

Bigelow, S.J. 2021. *Dark data discovery: How and where to find it*. https://www.techtarget.com/searchitoperations/tip/Dark-data-discovery-How-and-where-to-find-it (Accessed 11 September 2023).

Cafarella, M., Ilyas, I., Kornacker, M., Kraska, T. & Ré, C. 2016. Dark data: Are we solving the right problems? In 2016 IEEE 32nd International Conference on Data Engineering (ICDE), pp/ 1444-1445. https://ieeexplore.ieee.org/abstract/document/7498366 (Accessed 15 April 2023).

Chartered Governance Institute UK & Ireland. 2023. *The necessity of wellbeing as a part of health and safety governance*. https://www.cgi.org.uk/blog/wellbeing-health-safety-governance (Accessed 27 December 2023).

CIoD, Nigeria. 2022. *Increasing transparency and efficiency: The role of data governance in corporate governance*. https://www.iodnigeria.org/blog/ciod-weekly-1/increasing-transparency-and-efficiency-the-role-of-data-governance-in-corporate-governance-31 (Accessed 23 December 2023).

Corallo, A. Et al. 2021. Understanding and defining dark data for the manufacturing industry. *IEEE Transactions on Engineering Management* (Accessed 15 April 2023).

Corporate Governance Institute. 2020. *Enquire now introducing board members to data analysis and its importance*. https://www.thecorporategovernanceinstitute.com/insights/guides/how-to-use-data-to-make-better-decisions/ (Accessed 26 December 2023).

Crown Records Management. n.d. *Records management perspectives*. https://thevault.crownrms.com/download/262 (Accessed 27 May 2019).

Cubeware GmbH. 2021. *Dark data: Definition, types, examples, advantages, and disadvantages*. https://www.linkedin.com/pulse/dark-data-definition-types-examples-advantages-disadvantages- (Accessed 12 April 2023).

Dash, U.K. 2022. *How to leverage dark data for customer support*. https://fluentsupport.com/dark-data-for-customer-support/ (Accessed 16 September 2023).

DFIN. 2022. *What is dark data? How to manage and protect it*. https://www.dfinsolutions.com/knowledge-hub/thought-leadership/knowledge-resources/dark-data (Accessed 08 August 2023).

Donohue, J. 2022. *Good governance: 9 principles to set your organisation up for success.* https://www.diligent.com/resources/blog/what-constitutes-good-governance (Accessed 02 September 2023).

Firican, G. 2023. *What is dark data?* (a clear explanation). https://www.lightsondata.com/what-is-dark-data-a-clear-explanation/#t-1636448626038 (Accessed 22 July 2023).

Gartner Inc. 2023. *Dark data.* Available at https://www.gartner.com/en/information-technology/glossary/dark-data (Accessed 12 April 2023).

Hand, D.J. 2020. *Dark data: Why what you don't know matters*. https://press.princeton.edu/books/hardcover/9780691182377/dark-data (Accessed 07 August 2023).

Hanvey, L. 2023. *What is dark data and why does it matter to public agencies?* https://urbanlogiq.com/what-is-dark-data-and-why-does-it-matter-to-public-agencies/ (Accessed 07 September 2023).

Kamau, J. 2019, February 4. Detectives track Sh14.2 billion scam at power supply company. *Daily Nation*. https://www.nation.co.ke/news/Power-firm-caught-in-Sh14-billion-scam/1056-4965136-14ecjxfz/index.html (Accessed 03 December 2023).

KDNuggets. 2022. *The rise of dark data and how it can be harnessed*. https://www.kdnuggets.com/2016/03/rise-dark-data-how-harnessed.html (Accessed 29 October 2023).

Kenya Insights Team. 2022, February 10. *Scandal ridden KETRACO boss Fernades Barasa runs away with Sh800M power saga hanging on his neck.* https://kenyainsights.com/scandal-ridden-ketraco-boss-fernades-barasa-runs-away-with-sh800m-power-saga-hanging-on-his-neck/ (Accessed 07 June 2024).

Kizen. 2023. *What is dark data and why should you care?* https://kizen.com/content/data/what-is-dark-data/ (Accessed 13 September 2023).

Labovich, M. 2022. *How dark data can unlock real value by understanding risk.* https://www.pwc.com/us/en/tech-effect/ai-analytics/how-dark-data-can-unlock-real-value.html (Accessed 22 July 2023).

ManageEngine. 2023. *Dark data*. https://www.manageengine.com/data-security/what-is/dark-data.html (Accessed 07 August 2023).

Mann, T. 2022. *What is dark data and how to find it.* https://www.lepide.com/blog/what-is-dark-data-and-how-to-find-it/ (Accessed 15 July 2023).

Mosweu, O. & Rakemane, D. 2020. *The role of records management in ensuring good governance in Africa: Impediments and solutions.* https://www.ajol.info/index.php/jsasa/article/view/202314 (Accessed 27 December 2023).

Neha & Pahwa, P. 2020. Dark data analytics using blockchain technology. In Advances in data sciences, security and applications: Proceedings of ICDSSA 2019, pp. 467-474. Springer Singapore.

Odhiambo, M. 2022. MPs query power tender award to Spanish contractor. *The Star*. https://www.the-star.co.ke/news/2022-02-17-mps-query-power-tender-award-to-spanish-contractor/ (Accessed 07 June 2024).

Permutable.ai. 2023. *The transformative role of data and AI in corporate governance.* https://www.linkedin.com/pulse/transformative-role-data-ai-corporate-governance-permutableai (Accessed 27 December 2023).

Raghavan, S. 2023. *What is dark data and how can you make sense of it?* https://atlan.com/dark-data-101/ (Accessed 07 September 2023).

Rahmouni, M. 2023. *Dark data: Why what you don't know matters* by David J. Hand, Princeton, NJ: Princeton University Press, 2020, pp. xii+ 330. https://www.tandfonline.com/doi/full/10.1080/00401706.2022.2163804 (Accessed 12 July 2023).

Rao, V. 2018. *Extracting dark data*. https://developer.ibm.com/articles/ba-data-becomes-knowledge-3/ (Accessed 17 June 2024).

Sarkar, A. 2022. *How 'Dark data' is affecting the environment*. https://theswaddle.com/how-dark-data-is-affecting-the-environment/#:~:text=Dark%20data%20stored%20away%20in,chunk%20of%20internet%2Drelated%20emissions.&text=This%20requires%20significant%20attention%2C%20especially%20amidst%20the%20ongoing%20climate%20crisis (Accessed 08 August 2023).

Simplilearn. 2023. *What is data: Types of data, and how to analyze data?* https://www.simplilearn.com/what-is-data-article (Accessed 03 November 2023).

Singh, V. & Miller, M. 2024. *Democratize data access and data discovery: Empower your organisation to create and share data insights.* Blog. https://www.starburst.io/blog/democratize-data-access-and-discovery/ (Accessed 17 June 2024).

Smith, D. 2022. *Managing employee wellbeing is part and parcel of good governance*. https://www.enterprisecare.com.au/managing-employee-wellbeing-is-good-governance/ (Accessed 28 December 2023).

Sonawe, N. 2022. *Strategies to make the most out of dark data*. https://enterprisetalk.com/featured/strategies-to-make-the-most-out-of-dark-data/ (Accessed 10 August 2023).

Sonawe, N. 2023. *Risks of dark data*. https://enterprisetalk.com/featured/risks-of-dark-data/ (Accessed 09 August 2023).

Soni, P. 2022. *Dark data - meaning, present trends, pros, cons and risks.* https://www.analyticssteps.com/blogs/dark-data-meaning-present-trends-pros-cons-and-risks (Accessed 10 August 2023).

Splunk. 2019. *What is dark data?* https://www.splunk.com/en_us/data-insider/what-is-dark-data.html (Accessed 15 July 2023).

Srinivasa, R. 2022. *What is dark data and how can you make sense of it?* https://atlan.com/dark-data-101/ (Accessed 15 July 2023).

Steel, K. 2022. *What is dark data and why must you find it?* https://cloudsecurityalliance.org/blog/2022/03/11/what-is-dark-data-and-why-must-you-find-it/ (Accessed 07 August 2023).

Tittel, E. 2014. *The dangers of dark data and how to minimize your exposure.* https://www.cio.com/article/250529/the-dangers-of-dark-data-and-how-to-minimize-your-exposure.html (Accessed 13 August 2023).

UNESCO-IBE. 2023. *Concept of governance.* https://www.ibe.unesco.org/en/geqaf/technical-notes/concept-governance#:~:text=Governance%20has%20been%20defined%20to,%2C%20and%20broad%2Dbased%20participation (Accessed 02 September 2023).

United Nations. 2024. *About good governance – OHCHR and good governance.* https://www.ohchr.org/en/good-governance/about-good-governance (Accessed 20 January 2024).

Van Loon, R. 2023. *Dark data: What it is & how businesses should address it*. https://www.simplilearn.com/what-is-dark-data-article (Accessed 15 July 2023).

Wanyaga, F., Gikandi, J. & Ndirangu, P. 2017. *Dark data: Business analytical tools and facilities for illuminating dark data*. LAP LAMBERT Academic Publishing.

World Economic Forum. 2023. *What is 'Dark data' and how is it adding to all of our carbon footprints?* https://www.weforum.org/agenda/2022/10/dark-data-is-killing-the-planet-we-need-digital-decarbonisation/ (Accessed 22 July 2023).

Yackel, R. 2022. *What is dark data and how it causes data quality issues.* https://databand.ai/blog/what-is-dark-data-and-how-it-causes-data-quality-issues/ (Accessed 22 July 2023).

York, T. 2023. *Dark data: Discovery, uses & benefits of hidden data*. https://www.splunk.com/en_us/blog/learn/dark-data.html (Accessed 07 September 2023).