Journal of Social Research

P-ISSN:2827-9832
E-ISSN: 2828-335x
Vol. 5, No. 9, August 2026
http://ijsr.internationaljournallabs.com/index.php/ijsr

How Does IT Governance Regulate the Use of Big Data as a Supplement to Statistical Data?

Ridha Asih*, Febriliyan Samopa
Institut Teknologi Sepuluh Nopember, Indonesia
Email: ridhaasih@gmail.com*, iyan@its.ac.id

Keywords:
big data; Capability Level; COBIT 2019; IT governance.

ABSTRACT

This study aims to design an effective governance framework for the utilization of big data as a complement to official statistical data, while preventing new complexities in data management and utilization processes. The research employs a quantitative descriptive approach based on the COBIT 2019 framework, focusing on the governance objectives selected through the design factor process, namely EDM01 (Ensured Governance Framework Setting and Maintenance) and EDM05 (Ensured Stakeholder Engagement). The object of this research was the official statistics office in Balikpapan City, East Kalimantan, with the scope limited to governance policy design and excluding technical software or tool implementation. Data were collected through questionnaires based on the COBIT 2019 Process Assessment Model, supporting documents, and stakeholder discussions, then analyzed through design factor analysis, capability level assessment, and gap analysis. The results indicate that both EDM01 and EDM05 are currently at Capability Level 2 (Performed Process), with most activities classified as Partially Achieved and Largely Achieved, indicating that governance processes have been implemented but are not yet fully standardized and consistently managed. Based on these findings, the target state (To-Be) is set at Capability Level 3 (Defined). Recommendations include formulating Big Data governance policies, establishing governance structures, developing standard operating procedures, implementing periodic monitoring and evaluation, and strengthening stakeholder engagement and communication to support structured, integrated, and accountable big data governance.

INTRODUCTION

Big data is a term given to data sets that are so large and complex that they cannot be processed using conventional database management devices or other data processing applications (Maryanto, 2017). Big data is currently emerging as a new alternative source that complements statistical data. Based on its characteristics, Big Data was first introduced by Doug Laney (2001) through the concept of 3V (Volume, Velocity, Variety). As technology developed, the concept was developed into 5V with the addition of Veracity and Value, which are currently the most widely used characteristics in research and industrial practice.

Some of the methods that are often used today in pulling information from big data are the We Scraping and Crawling methods. Generally, this process is done by building a DOM (Document Object Model) tree structure first, then accessing essential data through that structure during the extraction process. However, depending on the complexity of the data structure in the DOM tree, the construction process of this tree can significantly increase computational time (Pamuji et al., n.d.; Sharma, 2024; Surahman et al., n.d.). Web scraping is also known by other terms such as web data extraction, web data scraping, web data collection, or screen scraping (Sultan & Abdullah, 2022).

Today's approach to producing quality data and information no longer places basic statistics and Big Data as two competing methodologies, but rather as two complementary instruments. Although it has great potential as a complement to data, both in terms of size and depth of data, bigdata is still unstructured and prone to bias, so adjustments are needed according to the needs in basic statistics that have been standardized and have a clear methodology (Holden, 2016; Kitchin, 2021; McCue & McCoy, 2017; Müller et al., 2016; Wenzel & Van Quaquebeke, 2018).

With the complexity of data and the diversity of sources used in the use of big data, structured and systematic governance is needed to ensure that the process runs consistently, efficiently, and responsibly. From the perspective of data governance frameworks such as COBIT 2019, this synergy emphasizes that adaptive governance is needed to ensure that the diversity of data formats (variety) and its flow rate (velocity) can be filtered using valid statistical principles, to be able to transform the raw data pile into accountable added value for the organization.

Ignoring governance in the use of Big Data poses critical risks that can be a huge burden on organizations. Regulations that run well can create added value/trust in basic statistical data, which has strict regulations and monitoring in its processing, thereby increasing the credibility of the institution in the eyes of the public (veracity risk). In addition, the absence of access control and operational standardization triggers the vulnerability of abuse of authority to fatal cybersecurity vulnerabilities to personal and confidential information protected by law (Elendu et al., 2024; Leverett et al., 2017; Obaidat et al., 2020; Qureshi & Koo, 2026; Yaqoob et al., 2019).

The term governance initially only referred to corporate governance or better known as corporate governance, which was designed so that company management could run professionally. The information technology function in the company is considered only as a very important supporting function but not as the main business activity (Cahaya et al., n.d.). *Governance* is a derivative of the word *government*, which means making policies that are in line with the desires/aspirations of the community or constituents (Handler & Lobba, 2005).

The COBIT best practice framework was developed and promoted to assist in the process of understanding, designing and implementing Enterprise Governance Information and Technology (EGIT). COBIT stands for Control Objective of Information and Related Technology, which is built on the integration of more than 25 years of development in this field, not only incorporating new insights from science, but also operationalizing these insights as a practice (ISACA, 2012).

The urgency of this research is underscored by the increasing use of big data as a complement to official statistics in Indonesia, particularly at the regional level. According to the National Statistics Office (BPS), the use of big data sources such as Mobile Positioning Data has been implemented since 2016 for measuring cross-border foreign visitors, with expansion to domestic tourism and metropolitan mobility in 2018. However, challenges regarding big data governance have become potential threats to data sustainability and the entire data provision process. The Balikpapan City statistics office, as the research locus, requires structured governance to ensure accountable and reliable big data utilization.

The novelty of this research lies in its comprehensive application of the COBIT 2019 framework through systematic design factor analysis, capability level assessment, and gap

analysis to produce specific governance recommendations for big data utilization in official statistics. This research provides empirical evidence of current capability levels (As-Is) and formulates targeted recommendations to achieve the desired capability level (To-Be), offering a replicable model for other statistics offices in Indonesia.

Through the IT governance approach using the COBIT 2019 framework, this research can harmonize the characteristics of Veracity and Value of Big Data through valid statistical principles, this research will not only produce a standardized and accountable governance design. In addition, it also makes a real contribution in mitigating various IT risks & legal regulations in optimizing the use of Big Data as a trusted official statistical complementary instrument.

METHOD

This research adopted a quantitative descriptive approach to design big data governance using the COBIT 2019 (Control Objectives for Information and Related Technologies) framework, focusing on the governance objectives selected through the design factor process (EDM01 and EDM05). The object of this research is the official statistics office in the city of Balikpapan, East Kalimantan, with the scope limited to the design of governance policy and excluding the technical aspects of software or tools implementation.

Data were collected through three complementary techniques. First, questionnaires developed based on the COBIT 2019 Process Assessment Model (PAM) were distributed to top management and staff who have authority and direct understanding of big data management, to measure the achievement level of each governance process activity. Second, these questionnaire results were reinforced through supporting documents and internal reports at the research locus, used to verify and validate respondents' answers against actual organizational conditions. Third, discussions with stakeholders were conducted during the design factor stage to confirm the organization's strategic priorities and governance scope.

Data analysis was carried out in three stages. First, a design factor analysis was performed to determine the organization's governance priority scope, based on ten COBIT 2019 design factors (e.g., enterprise strategy, enterprise goals, risk profile, IT-related issues, and compliance requirements). Second, a capability level assessment was conducted for each selected governance objective by calculating the percentage of achievement of process activities against four rating categories Not Achieved (0–15%), Partially Achieved (>15–50%), Largely Achieved (>50–85%), and Fully Achieved (>85–100%) in accordance with the COBIT 2019 Process Assessment Model. Third, a gap analysis was performed by comparing the current capability level (As-Is) with the intended target level (To-Be) to formulate governance improvement recommendations.

RESULTS AND DISCUSSION

The process of determining the organizational context and strategy is carried out by filling in factor design as a systematic step in the implementation of the Cobit 2019 framework (ISACA, 2012). The real condition of the organization is identified through various aspects ranging from strategy, risk profile, compliance with regulations, to technology adoption strategies that are relevant to the organizational context.

The assessment of design factors is carried out by referring to the standards contained in COBIT 2019, and applied in the context of organizational operations. The process of filling in this design factor involves parties who have authority and a deep understanding of data management, namely the organization's top leadership.

The organization's strategy does not focus on revenue or asset growth, but still supports the sustainability of intangible assets, such as software development and licensing, while prioritizing short-term cost efficiency. Therefore, the value of Growth and Cost Leadership is given an importance level of 1. Meanwhile, the organization continues to encourage Innovation/Differentiation through innovations that simplify business processes and improve the quality of service to the community. By 2025, more than 20 business process and service innovations have been produced.

The assessment of enterprise goals against the 13 calcifications of the Cobit 2019 framework represents the main goals to be achieved in designing a governance system in accordance with institutional needs.

Table 1. Enterprise Goals Design Factor Table

Value	Importance (1-5)	Baseline
EG01— Portfolio of competitive products and services	2	3
EG02— Managed business risk	2	3
EG03— Compliance with external laws and regulations	5	3
EG04— Quality of financial information	4	3
EG05— Customer-oriented service culture	5	3
EG06— Business-service continuity and availability	2	3
EG07— Quality of management information	5	3
EG08— Optimization of internal business process functionality	5	3
EG09— Optimization of business process costs	4	3
EG10— Staff skills, motivation and productivity	5	3
EG11— Compliance with internal policies	5	3
EG12— Managed digital transformation programs	5	3
EG13— Product and business innovation	2	3

Based on the organization's strategic plan document, compliance with regulations and legal compliance in the implementation of a reliable and reliable statistical system has an importance value 5. *Enterprise goals* also focus on service satisfaction and information management quality by optimizing every function in business processes.

This assessment was also carried out on the design of risk profile factors and IT-Related Issues, which received a scale of 3 assessment and became a serious issue, namely significant IT-related incidents, such as data loss, security breaches, project failures, or application errors, which are consistent with risk categories identified in other information security risk studies (Huang et al., 2024). Therefore, based on the results of the processing of all COBIT 2019 design factors, it can be concluded that the main priorities of organizational IT governance and management focus on managing organizational changes, managing programs and projects, and the availability and sustainability of IT services.

Based on the results of the Threat Landscape design factor questionnaire, the threat level faced by organizations is in the High and Normal categories with a proportion of 50%

each, indicating that the threat level is still moderate. In the design factor of Compliance Requirements, BPS Balikpapan City has a very high level of compliance (100%) considering its role as an official statistical organizer who must comply with regulations, national policies, and data management standards (BPS Balikpapan, 2025). In terms of the Role of IT, information technology plays a major role as a core operational (Factory) with a score of 5, as well as as a support for business processes (Support) and a support for strategic changes (Turnaround) with a score of 4, while the strategic role (Strategic) gets a score of 3.

In the IT Sourcing Model factor, organizations prioritize IT management internally (70%), supported by cloud utilization (30%), while outsourcing is not an option. For IT Implementation Methods, organizations are fully implementing the Agile approach (100%), while DevOps and Traditional have not been used. Meanwhile, in the Technology Adoption Strategy factor, organizations tend to play the role of Followers (80%), with a small percentage applying the First Mover approach (20%), and excluding Slow Adopters.

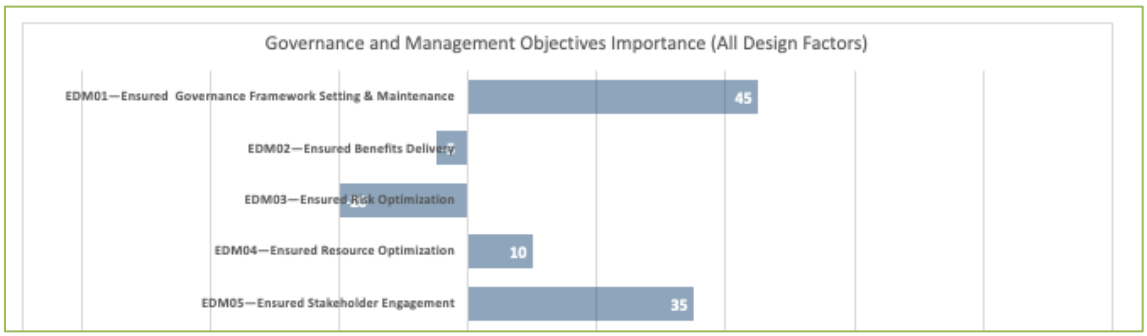

Figure 1. Conclusion of the whole design factor

Initial Scope of Governance

The determination of the governance scope is the initial stage in the implementation of the COBIT 2019 framework, carried out to identify which governance and management objectives are most relevant to the organization's priorities. Based on the results of the design factor processing, it was found that several objectives obtained a priority scale above 50, indicating that these objectives have a significant level of relevance to the organization's strategy, risk profile, and operational conditions. These objectives then become candidates for governance focus that are further confirmed through discussion with stakeholders, who have direct authority and understanding of the organization's data management needs.

Based on the results of this discussion, it was determined that the baseline to be achieved is a value of 35 for the governance domain and a value of 65 for the management domain, reflecting the organization's priority proportion between strategic oversight functions (governance) and operational implementation functions (management). Referring to this baseline determination, two objectives within the governance domain were selected to be assessed at the capability level, namely Evaluate, Direct and Monitor 01 (EDM01) Ensured Governance Framework Setting & Maintenance, and Evaluate, Direct and Monitor 05 (EDM05) Ensured Stakeholder Engagement. Both objectives were chosen because they are considered the most representative in reflecting the organization's governance needs in managing the use of big data as a complement to official statistical data.

EDM01 Ensured Governance Framework Setting & Maintenance

A core process in IT governance that focuses on providing a consistent, integrated, and aligned approach to governance as a whole. The goal is to ensure that every decision related to technology and information (I&T) is made in line with the organization's strategy and goals to achieve the desired value.

In practice, this domain demands an in-depth analysis of governance needs as well as a clear determination of responsibilities and authority, and transparent and effective oversight. Its activities include evaluating governance needs and design by involving stakeholders, providing direction through the establishment of clear principles, processes, and decision-making mechanisms, and monitoring governance effectiveness to ensure value creation for the organization.

EDM05 Ensured Stakeholder Engagement

This domain aims to ensure stakeholder engagement in supporting information and data technology strategies through effective and timely communication and reporting. This process helps organizations maintain alignment between IT goals and strategies with organizational goals, while also setting goals, performance indicators, and necessary corrective actions.

In managing the use of big data as a statistical reference, EDM05 consists of three main activities, namely: EDM05.01 evaluates the needs of stakeholder engagement as well as communication and reporting needs; EDM05.02 directs the implementation of communication, reporting, and validation mechanisms based on the results of risk analysis; and EDM05.03 monitors the effectiveness of stakeholder engagement and reporting as a basis for evaluation and governance improvement.

Capability level

The assessment of the *overall objective* almost meets the score above 50%, which indicates that the organization has carried out most of the organizational change management activities, but the implementation is still not completely consistent and is still Partially Achieved (P) and Largely Achieved(L). This condition shows that change management has been underway, but it still needs formal and continuous strengthening in order to reach a higher level of capability level. So, this domain is still at level 2.

Table 2. Capability Level Table

Objectives	Current Level (As-is)	Intended Level (to-be)	GAP
EDM01 Ensured Governance Framework Setting & Maintenance	2	3	1
EDM05 Ensured Stakeholder Engagement	2	3	1

Based on the results of the capability level assessment, it was obtained that all *objectives* are EDM01 and EDM05 at Capability Level 2 (*Performed Process*) in the current condition (*As-Is*). This condition shows that these processes have been implemented and can achieve their goals through the implementation of basic activities, but they are not yet fully standardized and managed consistently throughout the organization. This gap analysis refers to the definition *of Capability Level* 0–5, where Level 2 represents the processes that have been executed, and Level 3 indicates the processes that have been defined and standardized. Therefore, the results

of this gap analysis are the basis for the formulation of improvement recommendations that are directed at meeting the characteristics of the next capability level in stages in accordance with the COBIT 2019 framework (S. & Ajismanto, 2023).

Recommendations for the Governance Draft for the Utilization of Big Data *Objectives* EDM 01

To achieve the Fully Achieved condition at Level 2 as the foundation for Level 3, several improvement steps need to be taken, namely: the preparation of a governance policy for the use of Big Data that contains its objectives, scope, and management principles; the establishment of a governance structure that explains the roles, responsibilities, and authority of each unit; the preparation of Standard Operating Procedures (SOP) that govern the entire process of using Big Data from the planning stage to evaluation; The development of monitoring, evaluation, and reporting mechanisms that are carried out periodically as the basis for process control and improvement; as well as increased coordination and socialization to all stakeholders so that policies and procedures can be understood and implemented consistently in each work unit (Budiwiyono Lie et al., 2024).

Based on these recommendations, EDM01's capability enhancement not only focuses on improving activities that are not yet optimal, but also on establishing governance that is documented, standardized, and consistently implemented. As such, organizations have a stronger foundation for achieving Capability Level 3 (Defined), which is a condition when the entire governance process has been defined, documented, and executed uniformly across the organization.

Recommendations for the Governance Draft for the Utilization of Big Data *Objectives* EDM 05

To achieve the Fully Achieved condition at Level 2 as the basis for Level 3, several improvement efforts are needed, namely: (1) Grouping stakeholders based on their characteristics, information needs, and level of importance so that communication and reporting strategies are more effective; (2) Carry out regular communication and reporting on the use of Big Data as a statistical reference through established mechanisms; (3) Conduct routine evaluation of the performance of information technology utilization using measurable indicators as a basis for improving services and decision-making; (4) as well as compiling and implementing Standard Operating Procedures (SOP) that govern the process of communication, reporting, coordination, and use of information technology so that all activities can be carried out uniformly in each organizational unit, consistent with prior findings that structured, documented service and communication procedures are a key driver of higher IT governance maturity (Triyunsari & Sutabri, 2023).

Based on these recommendations, EDM05's capability enhancement is directed at the establishment of documented stakeholder engagement mechanisms, structured communication, and consistent and documented reporting. With the fulfillment of these aspects, the process of managing relationships with stakeholders will become more effective, transparent, and sustainable so that it can support the achievement of Capability Level 3 (Defined) in accordance with the principles of COBIT 2019.

CONCLUSION

The use of Big Data as an official statistical reference has great potential in improving the completeness, timeliness, and coverage of statistical data. The implementation of the COBIT 2019 framework through the objectives EDM01 and EDM05 shows that improving governance requires not only the formulation of clear policies and organizational structures, but also the strengthening of stakeholder engagement, communication, reporting, and documented monitoring and evaluation mechanisms. The recommendations produced, such as the preparation of Big Data governance policies, the determination of roles and responsibilities, the preparation of standard operating procedures, the implementation of periodic monitoring and evaluation, the grouping of stakeholders, and the strengthening of communication and reporting, are strategic steps in standardized management in producing better quality data, as well as increasing trust in the official statistics produced.

REFERENCES

Badan Pusat Statistik Kota Balikpapan. (2025). *Kota Balikpapan dalam angka 2025*.

Budiwiyono Lie, L., Samopa, F., & Ginardi, R. V. H. (2024). Developing maturity matrix: Measuring banking sector readiness in digital corporate governance. In *Digest of technical papers—IEEE International Conference on Consumer Electronics* (pp. 720–726). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/ISCT62336.2024.10791107

Cahaya, P. T., Rahmani, R., Kemuning, J. L., Komplek, B., & Rahman, A. (n.d.). *Tata kelola teknologi informasi*.

Elendu, C., Omeludike, E. K., Oloyede, P. O., Obidigbo, B. T., & Omeludike, J. C. (2024). Legal implications for clinicians in cybersecurity incidents: A review. *Medicine, 103*(39), e39887.

Handler, & Lobba. (2005). *[Informasi sumber belum lengkap]*.

Holden, G. (2016). Big data and R&D management: A new primer on big data offers insight into the basics of dealing with "uncomfortable data"—data that is too large or too unstructured to be accommodated by a firm's existing processes. *Research-Technology Management, 59*(5), 22–26.

Huang, S. Y., Wang, T., Huang, Y. T., & Yeh, T. N. (2024). Information security risk items and management practices for mobile payment using non-financial-institution service providers: An exploratory study. *International Journal of Accounting Information Systems, 53*, 100684. https://doi.org/10.1016/j.accinf.2024.100684

ISACA. (2012). *COBIT 2019 implementation guide: Implementing and optimizing an information and technology governance solution.*

Kitchin, R. (2021). *The data revolution: A critical analysis of big data, open data and data infrastructures*.

Laney, D. (2001). 3D data management: Controlling data volume, velocity, and variety. *META Group Research Note, 6*(70), 1.

Leverett, E., Clayton, R., & Anderson, R. (2017). Standardisation and certification of safety, security and privacy in the "Internet of Things." *Proceedings of the 16th Workshop on the Economics of Information Security (WEIS)*.

Maryanto, B. (2017). *Big data dan pemanfaatannya dalam berbagai sektor*.

McCue, M. E., & McCoy, A. M. (2017). The scope of big data in one medicine: Unprecedented opportunities and challenges. *Frontiers in Veterinary Science, 4*, 194.

Müller, O., Junglas, I., vom Brocke, J., & Debortoli, S. (2016). Utilizing big data analytics for information systems research: Challenges, promises and guidelines. *European Journal*

of Information Systems, 25(4), 289–302.
Obaidat, M. A., Obeidat, S., Holst, J., Al Hayajneh, A., & Brown, J. (2020). A comprehensive and systematic survey on the Internet of Things: Security and privacy challenges, security frameworks, enabling technologies, threats, vulnerabilities and countermeasures. *Computers, 9*(2), 44.
Pamuji, G. C., Putra, Y. H., & Wahdiniwaty, R. (n.d.). *Using COBIT-19 to create information assurance of COVID-19 application* (pp. 61–65).
Qureshi, R., & Koo, I. (2026). A comprehensive survey of cybersecurity threats and data privacy issues in healthcare systems. *Applied Sciences, 16*(3), 1511.
S., & Ajismanto, F. (2023). Implementation evaluation of information technology in the new normal era using COBIT 2019 method. *KnE Social Sciences*. https://doi.org/10.18502/kss.v8i9.13318
Sharma, G. (2024). Web crawling and scraping: A survey. In *Proceedings of the 2024 International Conference on Healthcare Innovations, Software and Engineering Technologies (HISET 2024)* (pp. 190–192). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/HISET61796.2024.00063
Sultan, N. A., & Abdullah, D. B. (2022). Scraping Google Scholar data using cloud computing techniques. In *2022 8th International Conference on Contemporary Information Technology and Mathematics (ICCITM 2022)* (pp. 14–19). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/ICCITM56309.2022.10032044
Surahman, A., Ferico Octaviansyah, A., & Darwis, D. (n.d.). *Teknologi web crawler sebagai alat pengembangan market segmentasi untuk mencapai keunggulan bersaing pada e-marketplace*.
Triyunsari, D., & Sutabri, T. (2023). Analisis tingkat kematangan manajemen layanan pegawai berbasis teknologi informasi menggunakan framework COBIT 5 pada SMA Negeri 19 Palembang. *Indonesian Journal of Multidisciplinary on Social and Technology, 1*(2), 146–153. https://doi.org/10.31004/ijmst.v1i2.141
Wenzel, R., & Van Quaquebeke, N. (2018). The double-edged sword of big data in organizational and management research: A review of opportunities and risks. *Organizational Research Methods, 21*(3), 548–591.
Yaqoob, T., Abbas, H., & Atiquzzaman, M. (2019). Security vulnerabilities, attacks, countermeasures, and regulations of networked medical devices—A review. *IEEE Communications Surveys & Tutorials, 21*(4), 3723–3768.