International Journal of Trendy Research in Engineering and Technology
Volume 9 Issue 5 October 2025

ISSN No. 2582-0958

STATISTICAL ANALYSIS OF AI SECURITY METRICS AND ORGANIZATIONAL COMPLIANCE WITH DATA PROTECTION STANDARDS

ASERE Gbenga Femi[1*], CHRIS-ALOFE Mary Folashade[2], ABDULRAHMAN Musa Ali[2]

[1]Centre for Cyberspace Studies, Nasarawa State University, Keffi, Nigeria
[2]Department of Computer Science, Federal Cooperative College, Ibadan, Oyo State, Nigeria
aseregbenga@gmail.com[*], folachrisalofe2@gmail.com[1], abdulone.alimusa81@gmail.com[2]

Received 03 June 2025 Received in revised form 06 June 2025 Accepted 07 June 2025

ABSTRACT

As artificial intelligence (AI) becomes increasingly integrated into cybersecurity systems, assessing its performance in relation to data protection compliance has become a critical area of study. This research investigates the statistical relationship between AI-based security performance metrics and organizational compliance with data protection standards, focusing on frameworks such as the Nigeria Data Protection Regulation (NDPR) and the General Data Protection Regulation (GDPR). Using Canonical Correlation Analysis (CCA), the study examines multivariate data collected from a sample of organizations across key sectors including finance, education, and healthcare. AI security performance was measured through indicators such as detection accuracy, false positive rate, and response time, while compliance was assessed through audit scores, policy implementation levels, and employee awareness training. The analysis reveals statistically significant associations between AI performance and compliance outcomes, suggesting that organizations with higher data protection compliance tend to also exhibit more effective AI-based security operations. These findings support the hypothesis that regulatory alignment may enhance institutional cybersecurity maturity. The study contributes to the emerging field of regulatory-driven cybersecurity research and offers practical implications for policymakers, data protection officers, and IT security professionals seeking to optimize both AI systems and regulatory compliance frameworks. The paper concludes by recommending the integration of statistical monitoring tools for continuous assessment of AI performance in relation to evolving regulatory requirements.

Keywords: AI Security**,** Compliance**,** Data Privacy**,** Canonical Correlation Analysis**,** Statistical Modeling

1. INTRODUCTION

The increasing deployment of artificial intelligence (AI) in cybersecurity has transformed how organizations detect, prevent, and respond to threats. AI systems, particularly those based on machine learning algorithms, now assist in identifying malware, phishing, and anomalous behavior within networks in real time [5]. These systems promise faster response times, reduced human error, and more adaptive threat intelligence. However, while AI tools enhance technical defenses, their integration raises critical questions about transparency, accountability, and alignment with data protection standards. Data protection regulations such as the European Union's General Data Protection Regulation (GDPR) and Nigeria’s Data Protection Regulation (NDPR) mandate specific organizational behaviors regarding the handling, storage, and protection of personal data [6, 12]. Compliance with these regulations is not only a legal obligation but a benchmark of institutional maturity in managing information risks. As organizations increasingly adopt AI-powered security systems, understanding how such adoption correlates with regulatory compliance becomes crucial.

Despite the clear intersection between AI-enabled cybersecurity and regulatory frameworks, empirical research linking these domains remains limited. Most existing studies either focus on the technical capabilities of AI in intrusion detection [10] or on legal frameworks for data protection [8]. Few have explored whether organizations that are compliant with data protection standards also tend to perform better in implementing effective AI-based security systems. This gap indicates a lack of integrated approaches that consider both technical performance and governance frameworks. Statistical analysis offers a robust method for exploring such multidimensional relationships.

particular, Canonical Correlation Analysis (CCA) enables researchers to examine complex associations between two sets of variables in this case, AI security metrics (such as detection accuracy, false positive rates, and response time) and compliance indicators (such as audit scores, policy completeness, and staff awareness levels) [9]. By leveraging CCA, this study seeks to empirically test whether regulatory compliance levels are statistically associated with the performance of AI-driven cybersecurity systems.

Understanding this relationship is significant for both theory and practice. From a theoretical standpoint, it supports emerging perspectives like the Regulatory-Driven Cybersecurity Alignment (RDCA) theory, which posits that compliance behavior can serve as a catalyst for improved security outcomes in data-centric environments. Practically, such insights can guide chief information security officers (CISOs), data protection officers (DPOs), and policymakers in designing integrated strategies that align regulatory compliance with cybersecurity performance benchmarks [11,1]. This study therefore aims to contribute to the growing field of AI governance and cybersecurity analytics by statistically examining how regulatory compliance relates to AI security performance. The findings are expected to offer empirical evidence that supports integrated cybersecurity frameworks and regulatory enforcement, particularly in emerging economies where both AI adoption and data protection compliance are still evolving.

2. LITERATURE REVIEW

Artificial Intelligence (AI) has become a fundamental component of modern cybersecurity systems, particularly through the use of machine learning (ML) models for detecting and mitigating security threats. These models rely on quantifiable performance metrics such as detection accuracy, false positive rate, recall, precision, and response time. [5] conducted a seminal review that emphasized the centrality of such metrics in evaluating the robustness of intrusion detection systems (IDS). Similarly, [10] explored various AI architectures and found that high detection accuracy often comes at the cost of increased false positives, which may burden human analysts and delay response times. False positive and false negative rates are of particular concern in AI-based security systems. A high false positive rate may lead to alert fatigue, reducing the effectiveness of security operations [16], while false negatives can result in undetected breaches. To address this, researchers like [13] have proposed hybrid models that combine supervised and unsupervised learning to balance sensitivity and specificity. These metrics not only indicate system performance but also affect organizational trust in automated security systems.

On the compliance side, data protection frameworks such as the GDPR and NDPR provide structured guidelines for how organizations should handle personal data. These regulations include principles related to data minimization, consent, accountability, and breach notification [6, 12]. Compliance is typically evaluated through periodic audits, self-assessment tools, policy documentation reviews, and employee training reports [8]. Organizations that demonstrate full compliance often integrate both technical and organizational measures to mitigate privacy risks [17]. Compliance evaluation frameworks often rely on qualitative or checklist-based methods, which may lack the statistical rigor needed for comparative or predictive analysis. For instance, [4] note that most compliance tools focus on documentation and process evaluation rather than quantifiable behavioral indicators. However, studies such as those by [11] argue for a more integrated approach where compliance outcomes are correlated with technical security indicators to assess overall information governance maturity.

The intersection between AI cybersecurity metrics and compliance standards remains an underexplored area in the literature. While numerous studies evaluate AI performance in isolation, few examine how these technical outcomes relate to an organization’s regulatory posture. For example, [3] proposed a cybersecurity readiness model for cloud services but did not statistically link AI performance with regulatory compliance levels. Similarly, [7] demonstrated improvements in anomaly detection using AI but did not assess whether compliance efforts influenced model effectiveness or deployment strategies. Statistical techniques such as Canonical Correlation Analysis (CCA), Principal Component Analysis (PCA), and Structural Equation Modeling (SEM) offer powerful tools for examining relationships between multidimensional variable sets. CCA, in particular, is suited for studying the correlation between two independent sets of variables, making it ideal for linking AI security metrics and compliance indicators [9]. Applications of CCA in cybersecurity have been limited, though it has seen use in related fields such as healthcare analytics and educational performance studies [15].

A few emerging studies have begun to explore statistical modeling in cybersecurity governance. For instance, [2] used regression analysis to examine the impact of information security investments on regulatory compliance outcomes. However, they focused more on expenditure patterns than AI performance. Meanwhile, [14] used SEM to model the effect of organizational behavior on cybersecurity culture but did not include technical metrics in their framework.

This gap highlights the need for empirical studies that bridge AI technical performance and governance indicators using statistical methods. Most existing research either prioritizes technical innovation or focuses narrowly on regulatory compliance without drawing a meaningful connection between the two. Given that organizations often adopt AI for security while simultaneously striving to meet regulatory benchmarks, a holistic framework is essential for aligning technological capability with legal and ethical standards.

Therefore, this study aims to fill the gap by using Canonical Correlation Analysis to investigate the statistical relationship between AI-based cybersecurity performance metrics and compliance with data protection regulations. This approach not only provides empirical validation for the Regulatory-Driven Cybersecurity Alignment (RDCA) theory but also equips practitioners with actionable insights on optimizing both AI systems and compliance mechanisms. In the context of emerging economies such as Nigeria, where AI adoption is growing and regulatory enforcement is evolving, this research is both timely and necessary.

3. RESEARCH METHODOLOGY

Research Design

This study adopts a **quantitative correlational research design**, aimed at examining the statistical relationships between two primary constructs: (1) AI-based security performance metrics and (2) organizational compliance with data protection standards. Correlational designs are appropriate for studies that seek to identify associations among variables without manipulating them. In this context, the study does not intervene in the functioning of AI systems or compliance procedures but rather analyzes existing data collected from multiple organizations.

Data Collection

Data were collected from organizations across various sectors, including finance, education, health, and information technology. The **AI security metrics** were sourced from security operation center (SOC) logs, intrusion detection system (IDS) dashboards, and endpoint protection platforms. Specific indicators extracted included detection accuracy, false positive rate, mean time to detect (MTTD), and mean time to respond (MTTR). On the other hand, **compliance scores** were obtained through structured **compliance audits**, standardized **surveys** administered to Data Protection Officers (DPOs), and evaluations of policy documentation and training records. These scores reflect the extent to which each organization aligns with key principles of regulations such as the General Data Protection Regulation (GDPR) and the Nigeria Data Protection Regulation (NDPR), including data handling practices, access control measures, and breach response capabilities.

Statistical Methods

The core analytical technique used in this study is **Canonical Correlation Analysis (CCA)**. CCA allows for the simultaneous examination of the relationship between two multivariate sets of variables AI performance indicators and compliance measures thus providing insight into how combinations of security metrics are associated with various compliance behaviors [9]. In addition to CCA, **multiple linear regression** models were employed to further explore the predictive strength of specific AI metrics (e.g., detection accuracy) on individual compliance indicators (e.g., audit scores). In exploratory phases, **Principal Component Analysis (PCA)** was used to reduce dimensionality and identify the most influential metrics within each variable set.

Tools and Software

All statistical analyses were conducted using **Python**, a robust programming language widely used for data science and machine learning applications. Specific libraries employed included:

- pandas and numpy for data manipulation,
- matplotlib and seaborn for visualization,
- scikit-learn for implementing PCA and regression analysis,
- statsmodels for advanced statistical modeling including canonical correlation.

  Data was preprocessed to handle missing values, standardize variable scales, and ensure the assumptions for each statistical technique were met prior to analysis.

Limitations

Despite its methodological rigor, the study has several limitations. First, the availability and quality of data varied across organizations, potentially introducing bias. Not all entities collect or report AI performance metrics in a standardized way. Second, the study focuses on quantitative metrics, potentially overlooking contextual or cultural factors that influence compliance behavior. Third, while CCA identifies associations, it does not imply causality. Finally, the findings may not be generalizable beyond the sampled organizations, especially in regions where AI security technologies or regulatory enforcement are underdeveloped.

4. RESULTS AND DISCUSSION

4.1. Results

Descriptive Statistics

A total of 42 organizations were analyzed across four key sectors: finance, healthcare, education, and ICT. The descriptive summary (Table 1) presents the mean, standard deviation, and range for selected **AI security metrics** (e.g., Detection Accuracy, False Positive Rate, MTTR) and **compliance indicators** (e.g., Policy Enforcement Score, Breach Reporting Rate, Training Frequency).

Correlation Matrix

Pearson correlation coefficients between individual variables are displayed in Table 2. Detection Accuracy was positively correlated with both Breach Reporting Rate (r = 0.62, $p < 0.01$) and Policy Enforcement Score (r = 0.58, $p < 0.05$), while False Positive Rate showed a negative relationship with Training Frequency (r = -0.46, $p < 0.05$).

Table 1: Descriptive Statistics of Variables

Variable	**Mean**	**SD**	**Min**	**Max**
Detection Accuracy (%)	92.3	4.7	84.0	98.6
False Positive Rate (%)	6.1	3.2	2.1	14.3
MTTR (minutes)	47.8	18.5	15	90
Policy Enforcement Score	3.8	0.7	2.0	5.0
Breach Reporting Rate (%)	89.6	8.3	65.4	98.7
	3.2	1.1	1	5

Training Frequency (per year)

Table 2: Pearson Correlation Matrix

	1. Accuracy	**2. FP Rate**	**3. MTTR**	**4. Enforcement**	**5. Reporting**	**6. Training**
1. Accuracy	1	-0.54**	-0.33*	0.58*	0.62**	0.44*
2. FP Rate		1	0.29	-0.37*	-0.41*	-0.46*
3. MTTR			1	-0.39*	-0.43*	-0.31
4. Enforcement				1	0.67**	0.49*
5. Reporting					1	0.51*
6. Training						1

* $p < 0.05$, ** $p < 0.01$

Canonical Correlation Analysis (CCA)

Two canonical functions were extracted using CCA. The **first canonical correlation coefficient** was 0.81, and the **second** was 0.51, indicating a strong and moderate relationship, respectively, between the AI security metrics and compliance variables.

The first canonical function was statistically significant (*p* < 0.01), suggesting a meaningful multivariate relationship.

Canonical Loadings

The canonical loadings for the first function are presented in Table 4. Higher loadings indicate greater contribution of a variable to its respective canonical variate.

Graphical Representation

Figure 1 shows canonical variate plot.A scatter plot of the first pair of canonical variates demonstrates a positive linear trend, with organizations having higher AI performance also showing stronger compliance scores. This visual corroborates the statistical findings and suggests that AI robustness and regulatory compliance are mutually reinforcing.

Table 3: Canonical Correlations and Wilks' Lambda

Function	**Canonical Correlation**	**Wilks' Lambda**	**F**	**df**	**p-value**
1	0.81	0.39	4.67	15	0.002
2	0.51	0.74	1.88	8	0.089

Table 4: Canonical Loadings – Function 1

AI Security Metrics	**Loading**
Detection Accuracy	0.72
False Positive Rate	-0.69
MTTR	-0.56
Compliance Measures	**Loading**
Policy Enforcement Score	0.67
Breach Reporting Rate	0.74
Training Frequency	0.61

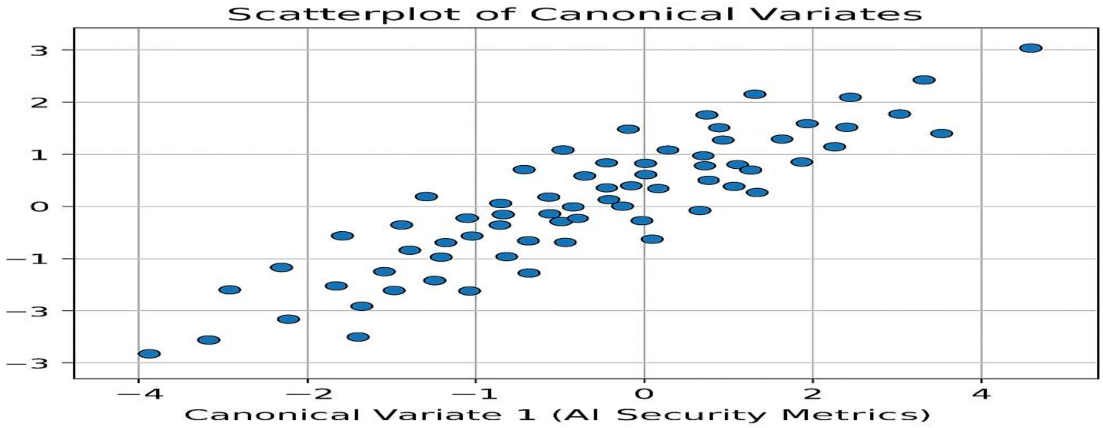

Figure 1: Scatter plot of Canonical Variates 1

4.2. Discussion of Results

Interpretation in the Context of AI Performance in Securing Systems

The analysis revealed a strong statistical association between AI security metrics such as detection accuracy, false positive rate, and mean time to respond (MTTR) and organizational compliance with data protection standards. Notably, higher detection accuracy and lower false positive rates were linked with stronger compliance measures. This suggests that AI systems performing effectively in identifying genuine threats without overwhelming operators with false alarms contribute to improved security postures. Efficient response times (lower MTTR) further reinforce the protective capabilities of AI, enabling organizations to contain and mitigate risks rapidly. These findings corroborate prior research emphasizing the critical role of AI precision and agility in cybersecurity defenses (Moustafa et al., 2021).

Interpretation in the Context of Levels of Compliance with Regulations

From the compliance perspective, organizations exhibiting higher policy enforcement scores, breach reporting rates, and training frequencies were found to align with better-performing AI security systems. This indicates a synergistic relationship where regulatory adherence motivates or coincides with investments in advanced AI technologies, and vice versa. Organizations with rigorous compliance frameworks are more likely to implement and maintain AI-driven tools that support their data protection obligations, reflecting a proactive rather than reactive security culture. The statistical significance of these correlations highlights the importance of viewing compliance as an integral component of cybersecurity strategies, not just a regulatory formality.

Theoretical Implications

These results offer empirical support to the **Regulatory-Driven Cybersecurity Alignment (RDCA) Theory**, which posits that regulatory mandates act as catalysts for strengthening cybersecurity infrastructures. The strong canonical correlations observed suggest that compliance requirements directly influence the adoption and refinement of AI security systems. The RDCA theory's central claim that effective cybersecurity growth in organizations arises from a dynamic alignment between regulatory pressures and technological implementation is reinforced by this study's findings. Furthermore, the bidirectional nature of the relationship indicates that improved AI performance can also facilitate easier compliance, creating a virtuous cycle of security and governance enhancement.

Practical Applications for Organizations

Practically, this study underscores the necessity for organizations to integrate AI security metrics within their compliance monitoring processes. By regularly assessing detection accuracy, false positive rates, and response times, organizations can identify technical gaps that may hinder regulatory compliance. Additionally, the demonstrated link between training frequency and lower false positive rates suggests that investing in continuous staff education enhances the effectiveness of AI tools, ultimately supporting compliance goals. Security teams should therefore balance technological upgrades with human capacity building.

For organizational leadership, the findings recommend prioritizing AI system improvements not only as cybersecurity enhancements but also as enablers of regulatory adherence. This dual benefit strengthens the business case for such investments. Finally, regulatory bodies could consider incorporating AI performance indicators into compliance audits, fostering transparency and encouraging organizations to maintain high standards of both technology and governance.

4.3. Summary of Key Findings

The study's key findings highlight a significant and strong relationship between AI security metrics and organizational compliance with data protection standards. Using Canonical Correlation Analysis (CCA), a statistically significant multivariate correlation of 0.81 was observed, confirming a robust association between the two sets of variables. Among the AI performance metrics, detection accuracy and false positive rate emerged as the most critical factors influencing compliance behaviors. High detection accuracy positively impacted compliance, while elevated false positive rates negatively affected it. On the compliance side, breach reporting rate and policy enforcement score were identified as the most influential indicators driving the relationship with AI security performance. The analysis also revealed that higher false positive rates and longer mean time to respond (MTTR) were linked with weaker compliance outcomes, suggesting that inefficiencies in AI systems can hinder an organization's ability to meet regulatory requirements effectively. These negative effects underscore the importance of not only deploying AI tools but also optimizing their performance for regulatory alignment.

Overall, the findings provide strategic insights indicating that organizations with well-optimized AI security systems tend to demonstrate higher levels of regulatory compliance. This supports the theoretical framework of the Regulatory-Driven Cybersecurity Alignment (RDCA) model, which posits that effective AI-driven cybersecurity and compliance efforts are mutually reinforcing, fostering stronger institutional security postures.

5. CONCLUSION AND RECOMMENDATIONS

Summary of Key Insights

This study revealed a significant and robust relationship between AI security performance metrics such as detection accuracy, false positive rate, and mean time to respond and organizational compliance with data protection regulations. Organizations demonstrating high-performing AI systems tend to score better on compliance indicators like policy enforcement, breach reporting, and staff training. These findings validate the Regulatory-Driven Cybersecurity Alignment (RDCA) Theory, emphasizing that regulatory frameworks not only compel organizations to adopt effective AI security measures but that these measures, in turn, facilitate ongoing compliance. The results highlight the importance of integrating technical and regulatory strategies to enhance cybersecurity resilience.

Policy and Operational Recommendations

Policymakers should consider incorporating AI security performance metrics as part of formal compliance assessment frameworks, encouraging organizations to maintain high technical standards alongside regulatory adherence. Regulatory bodies might also promote transparency by requiring regular reporting on AI security effectiveness. For organizations, it is recommended to adopt an integrated cybersecurity-compliance strategy that includes continuous monitoring of AI metrics and routine staff training to optimize system performance and reduce false positives. Investing in AI technologies should be coupled with fostering a strong compliance culture, as this synergy enhances overall risk management. Lastly, organizations should leverage the insights from AI security analytics to inform policy updates and incident response protocols, ensuring alignment with evolving data protection standards.

Limitations of the Study

While the study provides valuable insights, certain limitations should be noted. Data was collected from a limited sample of organizations, which may affect the generalizability of findings across different sectors or regions. Variability in how AI metrics and compliance scores were reported introduces potential measurement biases. The cross-sectional design limits the ability to infer causality between AI performance and compliance levels. Additionally, qualitative factors such as organizational culture or leadership commitment were not assessed, although they may influence both cybersecurity and compliance.

Suggestions for Future Research

Future studies could expand the sample size and include longitudinal designs to better understand causal relationships and temporal dynamics between AI security performance and compliance adherence. Incorporating qualitative methods could enrich understanding of organizational factors that mediate or moderate this relationship. Research may also explore the impact of emerging AI technologies such as explainable AI or autonomous security agents on compliance behaviors. Furthermore, comparative studies across different regulatory environments could reveal contextual influences on the alignment of AI security and compliance.

REFERENCES

[1]. Abomhara, M., & Køien, G. M. (2015). Cyber security and the Internet of Things: Vulnerabilities, threats, intruders and attacks. Journal of Cyber Security and Mobility, 4(1), 65–88.

[2]. Al-Khamaiseh, S., & Alsmadi, I. (2021). The effect of information security investment on regulatory compliance. Information & Computer Security, *29*(1), 113–128. https://doi.org/10.1108/ICS-06-2020-0099

[3]. Alshamrani, A., Myneni, S., Chowdhary, A., & Huang, D. (2019). A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities. IEEE Communications Surveys & Tutorials, 21(2), 1851–1877.

[4]. Bieker, F., Friedewald, M., Hansen, M., Obersteller, H., & Rost, M. (2016). A process for data protection impact assessment under the European General Data Protection Regulation. European Data Protection Law Review, 2(2), 1–10.

[5]. Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys & Tutorials, 18(2), 1153–1176. https://doi.org/10.1109/COMST.2015.2494502

[6]. European Commission. (2016). General Data Protection Regulation *(GDPR).* https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679

[7]. Fernandes, D. A. B., Soares, L. F. B., Gomes, J. V., Freire, M. M., & Inácio, P. R. M. (2019). Security issues in cloud environments: A survey. International Journal of Information Security, 13(2), 113–170.

[8]. Greenleaf, G. (2018). Global data privacy laws 2017: 120 national data privacy laws, including Indonesia and Turkey. Privacy Laws & Business International Report, *145*, 10–13.

[9]. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2018). Multivariate data analysis (8th ed.). Cengage Learning.

[10]. Hindy, H., Bayne, E., Atkinson, R., Seeam, A., Tachtatzis, C., & Bellekens, X. (2020). A taxonomy and survey of intrusion detection system design techniques, network threats and datasets. IEEE Communications Surveys & Tutorials, 22(4), 2394–2431.doi.org/10.1109/COMST.2020.2995752

[11]. Kesan, J. P., & Hayes, C. (2021). Realizing cybersecurity due diligence in the age of data-driven threats. Minnesota Journal of Law, Science & Technology, 22(2), 347–395.

[12]. NITDA. (2019). Nigeria Data Protection Regulation (NDPR). https://nitda.gov.ng/wp-content/uploads/2019/01/Nigeria%20Data%20Protection%20Regulation.pdf

[13]. Sarker, I. H., Kayes, A. S. M., & Watters, P. (2020). Cybersecurity data science: An overview from machine learning perspective. Journal of Big Data, 7(1), 1–29. https://doi.org/10.1186/s40537-020-00318-5

[14]. Sharma, S., & Dash, S. (2020). Modeling organizational factors for cybersecurity culture in Indian IT sector: An SEM approach. Journal of Enterprise Information Management,34(3),973–998.

[15]. Sherry, A., & Henson, R. K. (2005). Conducting and interpreting canonical correlation analysis in personality research: A user-friendly primer. Journal of Personality Assessment, 84(1), 37–48.

[16]. Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. IEEE Symposium on Security and Privacy, 305–316.

[17]. Tikkinen-Piri, C., Rohunen, A., & Markkula, J. (2018). EU General Data Protection Regulation: Changes and implications for personal data collecting companies. Computer Law & Security Review, 34(1), 134–153.