

Zero Trust–Based Cybersecurity Strategies for Protecting Big Data in Business Intelligence Systems

Mehdi Ghazanfari

Student, Enghelab-e Eslami Technical College, Technical and Vocational University (TVU), Tehran,
Iran

m.ghazanfari1384@gmail.com

Amirhasan Keshavarz Mohammadian

Student, Enghelab-e Eslami Technical College, Technical and Vocational University (TVU), Tehran,
Iran

amirhasankeshavarz9@gmail.com

Mehrdad Hamidzade

PhD in IT Management, Faculty Member, Department of Computer Science, Enghelab-e Eslami
Technical College, Technical and Vocational University (TVU), Tehran, Iran

en.hamidzade@gmail.com

Amin Yousefli

M.Sc. in Secure Computing, Faculty Member, Department of Computer Science, Enghelab-e Eslami
Technical College, Technical and Vocational University (TVU), Tehran, Iran

amin.yousefli@pardisiau.ac.ir

Abstract

Business intelligence (BI) platforms rely on large-scale, heterogeneous and fast-moving datasets to support strategic and operational decisions. Integrating big data and advanced analytics increases competitive advantage but also expands the attack surface and amplifies the impact of security incidents. Traditional perimeter-based security architectures are poorly suited to multi-cloud, data-fabric and remote-work environments in which data, users and workloads are highly distributed.

This paper proposes a Zero Trust-based, data-centric cybersecurity framework for protecting big data throughout the BI lifecycle. Drawing on recent work on big data security and privacy, cryptographic techniques for AI security, Zero Trust Architecture (ZTA) and cybersecurity's role in organisational performance,

the paper makes three contributions. First, it develops a BI-oriented threat model and lifecycle matrix that maps confidentiality, integrity, availability, privacy, AI-specific and insider threats to each stage of the big data lifecycle (ingestion, storage, processing, access, sharing, disposal). Second, it defines a Zero Trust BI reference architecture (ZT-BI) and a maturity roadmap across five capability domains—Permission, Infrastructure, Processes, Intelligence and People—tailored to BI environments and workloads.

Third, it proposes design patterns that show how to phase cryptography and privacy-preserving techniques in ways that balance strong security with BI performance and profitability requirements.

The framework integrates data governance, pervasive encryption and key management, fine-grained access control, Zero Trust network and workload segmentation, AI-driven security analytics, privacy-preserving analytics and integrity-assurance mechanisms such as tamper-evident logging. By aligning these measures with business and regulatory constraints, organisations can mitigate risk, improve the trustworthiness of analytics and support operational efficiency and profitability in data-driven decision-making.

Keywords: Big Data, Business Intelligence, Cybersecurity, Zero Trust Architecture, Privacy-Preserving Analytics

1. Introduction

Big data analytics has become central to modern business intelligence (BI). Organisations collect and integrate transactional records, clickstreams, IoT sensor outputs, social media, log data and external feeds to create data warehouses, data lakes and lakehouses. BI platforms use these assets for real-time dashboards, self-service reporting, predictive models and prescriptive optimisation across domains such as finance, healthcare, logistics and public services. (Hussain & Hajjar, 2024)

However, the same properties that make big data valuable—volume, variety, velocity and the ability to link previously separate datasets—also create security and privacy challenges. (Kantarcioglu & Ferrari, 2019) Data is ingested from numerous sources, often via APIs and streaming pipelines, stored in distributed cloud and on-premises infrastructures, and accessed by diverse stakeholders including data engineers, analysts, data scientists, applications and external partners. Each stage introduces vulnerabilities: compromised devices, insecure APIs, misconfigured storage, over-privileged accounts, weak monitoring and opaque machine learning models susceptible to manipulation or theft. (Hussain & Hajjar, 2024), (Kantarcioglu & Ferrari, 2019)

Big data also has macro-scale implications. Stephan argues that big data ecosystems should be understood as emergent systems whose capabilities—mass surveillance, manipulative targeting, AI-driven influence and sophisticated cyber operations—pose national security risks that current legal regimes struggle to address. (Stephan, 2025) Big data enables both offensive and defensive cyber operations, and its compromise or misuse can impact critical infrastructure and democratic processes.

Traditional perimeter-centric security models assume that threats originate outside the network and that entities inside the perimeter are largely trustworthy. In an era of cloud computing, remote work, mobile access and extensive third-party integrations, this assumption fails. (Nair, 2021) Zero Trust Architecture (ZTA) instead operates on the principle of “never trust, always verify”: every access request by a user, device or workload is authenticated, authorised and continuously evaluated based on context. (Mushtaq, Mohsin, & Mushtaq, 2025)

In parallel, cryptographic techniques such as homomorphic encryption, secure multiparty computation (SMC), lightweight cryptography and quantum-resistant schemes are being developed and deployed to protect data and models in AI systems. A recent bibliometric review finds a sharp rise in research on cryptographic techniques in AI security between 2020 and 2024, highlighting homomorphic encryption, SMC and blockchain as key technologies. (Taherdoost, Le, & Slimani, 2025) Yet these techniques introduce performance overheads that must be reconciled with BI latency and throughput requirements.

Cybersecurity is also an economic concern. A recent cross-sectional study of Nigerian deposit money banks shows that cybersecurity capabilities significantly predict both operational efficiency and profitability, supporting the view that security can be a dynamic capability that enhances performance rather than a pure cost. (Egerson, Williams, Aribigbola, Okafor, & Olaleye, 2024) Similar arguments are made in wider analyses of AI adoption, which stress that security, bias mitigation and data quality are prerequisites for sustainable value creation from AI. (Williams, 2024)

In this context, the central question of this paper is: How can organisations design and implement Zero Trust-based cybersecurity strategies to protect big data assets in BI systems, while sustaining operational efficiency and profitability?

1.1 Problem statement

Big-data-driven BI systems aggregate sensitive personal, financial and operational data at scale. Breaches can lead to financial losses, regulatory sanctions, reputational damage and national-level risks. (Stephan, 2025), (Kantarcioglu & Ferrari, 2019) At the same time, pressure to increase data-driven innovation and share data across organisational and national boundaries continues to grow. The challenge is to reconcile these pressures with robust security and privacy.

1.2 Objectives

The objective of this paper is to develop a BI-specific Zero Trust cybersecurity framework that:

- Explicitly models threats across the big data lifecycle in BI environments.
- Provides Zero Trust-aligned architectural patterns and controls for each lifecycle stage.
- Offers a maturity roadmap and metrics that organisations can use to plan and evaluate BI security improvements.
- Articulates how these measures can support, rather than undermine, BI-enabled efficiency and profitability. (Egerson, Williams, Aribigbola, Okafor, & Olaleye, 2024)

1.3 Research gap and contributions

Existing work falls into four partially overlapping streams. First, big data security and privacy research identifies challenges across storing, querying, linking, sharing and analysing big data, but largely at a technology-agnostic level rather than specifically for BI platforms. (Kantarcioglu & Ferrari, 2019) Second, BI-oriented treatments of cybersecurity and big data give high-level overviews of threats and countermeasures but rarely formalise a lifecycle-oriented threat model or Zero Trust architecture. (Hussain & Hajjar, 2024) Third, cryptographic and privacy-preserving techniques for AI and big data are well documented, yet their integration into BI stacks and performance trade-offs remain under-explored. (Taherdoost, Le, & Slimani, 2025) Fourth, Zero Trust architectures and maturity models such as ZeTuMM are mainly domain-agnostic and provide limited guidance on BI-specific concerns like multi-cloud data fabrics, analytics tooling and AI workloads. (Mushtaq, Mohsin, & Mushtaq, 2025) (Modderkolk, 2018)

This paper addresses these gaps through three main contributions:

1. BI-oriented threat model and lifecycle mapping. We construct an explicit threat model for big-data-driven BI and map it to the full data lifecycle (ingestion → storage → processing → access → sharing → disposal). The resulting matrix links major threat categories to representative attack vectors at each lifecycle stage.

2. Zero Trust BI reference architecture and maturity roadmap. Building on ZTA principles and ZeTuMM, we propose a BI-specific Zero Trust reference architecture (ZT-BI) and a four-level maturity roadmap across five capability domains (Permission, Infrastructure, Processes, Intelligence, People), with BI-centric examples and metrics. (Modderkolk, 2018), (Mushtaq, Mohsin, & Mushtaq, 2025)
3. Performance-aware design patterns. We propose pragmatic patterns for selectively applying advanced cryptography and privacy-preserving techniques where they add most value, enabling organisations to balance Zero Trust aspirations with BI performance and profitability goals. (Taherdoost, Le, & Slimani, 2025), (Egerson, Williams, Aribigbola, Okafor, & Olaleye, 2024)

Table 1 positions this work relative to representative prior studies.

Prior work	Focus	Gap for BI big-data security	How this paper addresses it
Kantarcioglu & Ferrari (2019) – big data, security & privacy	Research challenges across storing, querying, linking, sharing and analysing big data; introduces encryption, TEEs, access control, differential privacy.	Technology-centric; limited BI lifecycle view or organisational roadmap.	Formalises a BI-specific lifecycle and threat matrix, mapping controls to each stage.
Hussain & Hajjar (2024) – cybersecurity and big data analytics in BI	Narrative overview of threats (breaches, insider attacks, fraud) and controls (encryption, AI detection, blockchain, governance.)	Lacks structured lifecycle mapping and Zero Trust architecture.	Provides a lifecycle-linked threat model and a ZT-aligned BI architecture.
Taherdoost et al. (2025) – cryptographic techniques in AI security	Bibliometric review of cryptographic methods (HE, SMC, quantum, blockchain) for AI security.	Does not address how to embed these techniques in BI workloads or handle performance trade-offs.	Derives BI-oriented design patterns showing where advanced cryptography is most appropriate.
Mushtaq et al. (2025); Modderkolk (2018); Nair (Zero Trust Paper)	ZTA across domains and a general Zero Trust maturity model; detailed comparison of ZTA implementations and challenges.	Domain-agnostic; limited focus on BI data fabrics, analytics tools and AI-driven BI workloads.	Specialises ZTA and ZeTuMM concepts to BI, with BI-specific domains, examples and metrics.
Egerson et al. (2024); Williams (2024)	Empirical link between cybersecurity and efficiency/profitability; broader view of AI and ethics.	No concrete blueprint for Zero Trust big-data BI; performance implications mostly qualitative.	Integrates performance-aware patterns and maturity metrics directly into the ZT-BI roadmap.

2. Background and Related Work

2.1 Big Data and Business Intelligence Architectures

BI environments typically combine:

- Data ingestion layers that collect, cleanse and validate batch and streaming data from operational systems, IoT devices and external sources.
- Storage layers comprising data warehouses, data lakes and lakehouses across multi-cloud and hybrid infrastructures.
- Processing and analytics layers including ETL/ELT tools, distributed query engines, ML platforms and notebooks.
- Access and presentation layers such as dashboards, reports, APIs and embedded analytics. Each layer may span multiple cloud providers and on-premises environments. This distributed and multi-tenant nature complicates the enforcement of consistent security and privacy controls.

Hussain and Hajjar emphasise that these distributed, multi-technology ecosystems complicate consistent security enforcement and introduce vulnerabilities at each integration point. (Hussain & Hajjar, 2024)

2.2 Threat Landscape for Big-Data-Driven BI

The threat landscape can be framed around confidentiality, integrity, availability and privacy, expanded with AI-specific and insider dimensions: (Kantarcioglu & Ferrari, 2019), (Hussain & Hajjar, 2024)

- Confidentiality threats: data breaches from misconfigured storage, weak or missing encryption, compromised credentials, insecure APIs and over-permissive BI access.
- Integrity threats: data tampering, dataset poisoning and manipulation of training data or models, leading to misleading analytics and decisions. (Kantarcioglu & Ferrari, 2019)
- Availability threats: ransomware and denial-of-service attacks against data pipelines, storage clusters or BI portals. (Nair, 2021)
- Privacy threats: large-scale collection and linkage of personal data allows re-identification, even when data is nominally anonymised. Differential privacy research has shown that naive anonymisation fails when adversaries have auxiliary information. (Kantarcioglu & Ferrari, 2019)
- AI-specific threats: model inversion, membership inference, adversarial examples and model extraction. (Taherdoost, Le, & Slimani, 2025)
- Insider threats: malicious or negligent insiders—administrators, developers, analysts—misusing privileged access or bypassing policy. (Hussain & Hajjar, 2024)

At a national-security level, Stephan shows how big data enables advanced cyber operations, influence campaigns and cross-domain intelligence, and how its compromise can undermine state capabilities. (Stephan, 2025)

2.3 Security, Privacy and Cryptography for Big Data

Kantarcioglu and Ferrari summarise research challenges along the big data pipeline: encrypted storage and querying, privacy-preserving record linkage, risk-aware data sharing, and privacy-aware analytics. (Kantarcioglu & Ferrari, 2019) Encrypted storage schemes—from searchable encryption to homomorphic encryption—permit queries over ciphertext but can leak access patterns and struggle with scale. Trusted

Execution Environments (TEEs) provide hardware-backed secure enclaves for processing encrypted data, but side-channel attacks remain a concern.

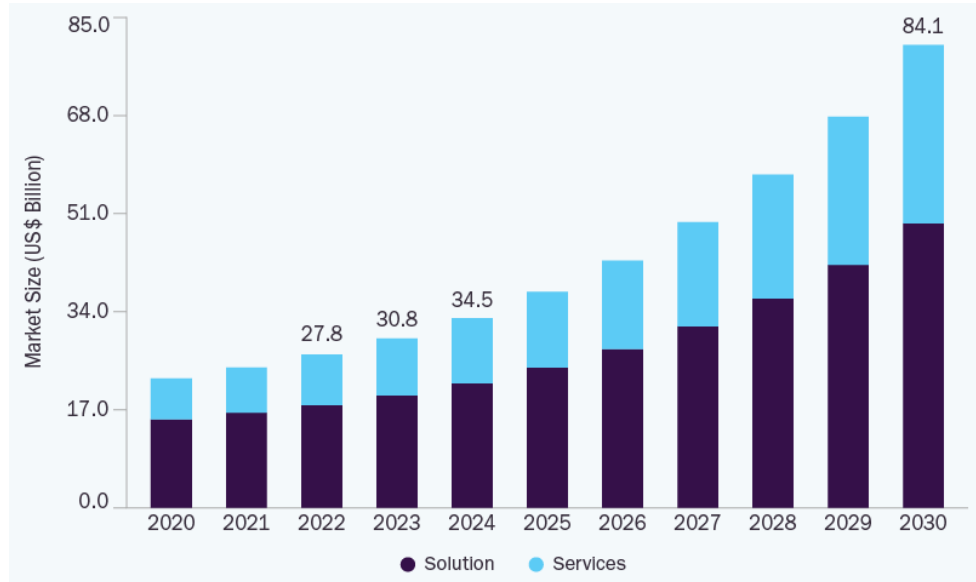


Figure 1 Zero Trust Architecture in analytics from 2020 to 2030

For data sharing and integration, private record linkage and differential privacy provide formal guarantees but can reduce utility or be difficult to scale across many parties and datasets. Practical risk-aware anonymisation tools offer more usable but less formally guaranteed techniques. (Kantarcioglu & Ferrari, 2019) A bibliometric review of AI security research shows accelerating interest in homomorphic encryption, SMC, quantum-resistant cryptography and blockchain for securing AI models and data, particularly in high-stakes domains. (Taherdoost, Le, & Slimani, 2025)

2.4 Zero Trust Architecture and Maturity Models

Zero Trust Architecture eliminates implicit trust based on network location, requiring continuous verification of user, device and workload identities and enforcing least-privilege access. NIST SP 800-207 and subsequent work describe key components: policy decision points (PDPs), policy enforcement points (PEPs), identity providers, trust engines and continuous monitoring. (Mushtaq, Mohsin, & Mushtaq, 2025)

Mushtaq et al. systematically review 74 ZTA implementations across domains including cloud, IoT, healthcare, industrial control, AI and big data, concluding that authentication, authorisation and access control are commonly addressed, while encryption, auditing and environmental perception are less mature. (Mushtaq, Mohsin, & Mushtaq, 2025)

Their paper includes a conceptual ZTA model (Figure 4 **Illustration of the proposed Zero Trust Business Intelligence (ZT-BI) architecture and its integrated components.**) that shows how user/device requests

are evaluated by PDPs using context signals and cryptographic channels before reaching protected resources.

Nair contrasts perimeter and Zero Trust models, demonstrating weaknesses of the former such as lack of intra-zone inspection, single points of failure and “phoning home” paths by which malware bypasses inbound controls via outbound traffic. (Nair, 2021)

Modderkolk’s ZeTuMM thesis proposes a Zero Trust Maturity Model that organises cybersecurity capabilities into five focus-area groups—Permission, Infrastructure, Processes, Intelligence and People—each containing capabilities with graded maturity levels. (Modderkolk, 2018) This provides a template for structuring a BI-specific maturity framework.

2.5 Cybersecurity, Operational Efficiency and Profitability

Egerson et al. empirically examine the effect of cybersecurity on operational efficiency and profitability in Nigerian deposit money banks. Using survey data and structural equation modelling, they show that cybersecurity capabilities significantly predict both efficiency and profitability, and recommend continuous investment in advanced cybersecurity technologies and skills. (Egerson, Williams, Aribigbola, Okafor, & Olaleye, 2024)

More broadly, Williams’ review of AI adoption across industries highlights security, ethical and data-quality concerns as critical to unlocking AI’s potential while managing risk. (Williams, 2024) Together, these works support the premise that robust cybersecurity is a driver of sustainable digital performance.

3. Research Approach and Conceptual Framework

This paper uses an integrative literature review and conceptual synthesis approach grounded in the sources above.

3.1 Literature search and selection

We conducted a structured search in Scopus, Web of Science, IEEE Xplore, ACM Digital Library and ScienceDirect, focusing on works that address big data or BI security, Zero Trust, cryptographic techniques for AI/big data, and the relationship between cybersecurity and organisational performance. Search queries included combinations such as:

- “big data” OR “data lake” OR “data warehouse” AND “security” OR “privacy” AND “business intelligence” OR “BI”;
- “Zero Trust” OR “Zero Trust Architecture” AND cloud OR “big data” OR “analytics”;
- “homomorphic encryption” OR “secure multiparty computation” OR cryptograph* AND “AI security” OR “machine learning security”;
- “cybersecurity” AND “operational efficiency” OR profitability OR performance.

We restricted to peer-reviewed articles, conference papers and theses in English, 2010–2025, reflecting the maturation period of big-data BI platforms and ZTA. Inclusion criteria were:

- Explicit discussion of big data or BI security, Zero Trust, or cryptographic techniques for analytics/AI;
- Relevance to organisational contexts (not just theoretical cryptography).

We excluded non-scholarly reports, opinion pieces and purely domain-specific case studies without generalisable security insight. From the resulting corpus, we grouped key works into four clusters (Table 2 *Main literature clusters informing the framework*).

Table 2 Main literature clusters informing the framework

Cluster	Typical keywords	Role in our framework	Representative sources
Big data & BI security/privacy	big data security, privacy, data governance, BI security	Defines core threat landscape and lifecycle stages for BI.	Kantarcioglu & Ferrari (2019); Hussain & Hajjar (2024)
Cryptographic & privacy-preserving techniques	homomorphic encryption, SMC, TEEs, differential privacy	Provides building blocks for data-centric, privacy-preserving controls.	Taherdoost et al. (2025); Kantarcioglu & Ferrari (2019)
Zero Trust architectures & maturity	Zero Trust, ZTA, micro-segmentation, maturity model	Informs ZT-BI reference architecture and maturity roadmap.	Mushtaq et al. (2025); Modderkolk (2018); Nair (Zero Trust Paper)
Cybersecurity & performance	cybersecurity capabilities, efficiency, profitability	Underpins the link between security, BI reliability and business outcomes.	Egerson et al. (2024); Williams (2024)

3.2 Big data lifecycle and threat matrix for BI

We conceptualise a BI big data lifecycle with six stages:

1. **Ingestion** – collection of raw data from operational systems, IoT, files and external feeds;
2. **Storage** – persistence in data lakes/warehouses, including raw, curated and sandbox zones;
3. **Processing** – ETL/ELT, feature engineering, ML training and inference;
4. **Access** – interactive querying, dashboards, notebooks and APIs;
5. **Sharing** – exchange of data, models or reports with external parties or the public;

6. **Disposal** – retention management, archival and secure deletion.

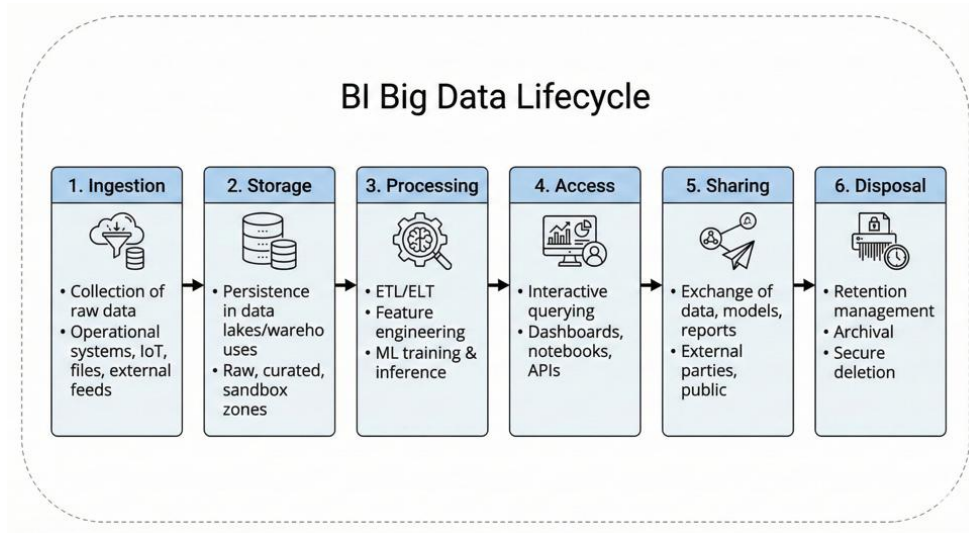


Figure 2 A diagram illustrating the six sequential stages of the BI Big Data Lifecycle, from data ingestion to disposal.

Table 3 *Threats mapped to the BI big data lifecycle* maps major threat categories to these stages with representative attack vectors. This matrix is subsequently used to link each proposed control to the cells it mitigates.

3.3 Conceptual framework

Our conceptual framework is layered:

- **Data governance layer:** policies for ownership, classification, retention, access and sharing.
- **Data-centric controls:** encryption, key management, fine-grained access control, privacy-preserving analytics and integrity mechanisms.
- **Zero Trust architecture layer:** micro-segmented networks and workloads, identity-aware proxies, PDP/PEPs and continuous trust evaluation.
- **Security analytics layer:** AI-driven monitoring, anomaly detection and automated response.
- **People and process layer:** governance bodies, incident response, training and culture.

Sections 4–7 instantiate this framework into concrete strategies, an architecture and a maturity roadmap.



ICAICS

<https://icaics.ir>
info@icaics.ir

اولین کنفرانس بین‌المللی هوش مصنوعی
و علوم کامپیوتری نو ظهور: از الگوریتم تا آینده‌نگری

**First International Conference on Artificial Intelligence
and Emerging Computer Science: From Algorithm to Foresight**

March 17, 2026-GEORGIA

۲۶ اسفند ماه ۱۴۰۴ - گرجستان

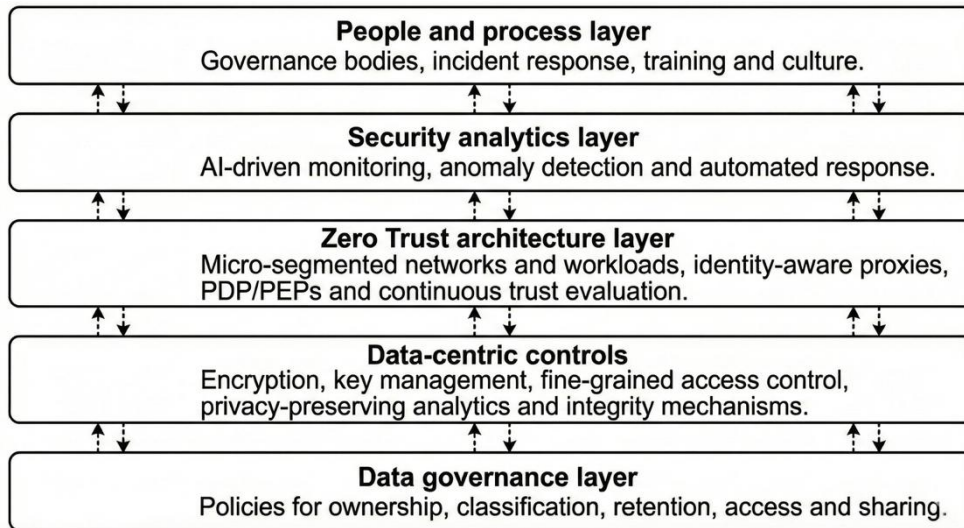


Figure 3 A diagram illustrating an interconnected five-layer framework for comprehensive security and data governance.

Table 3 Threats mapped to the BI big data lifecycle

Threat category	Ingestion	Storage	Processing	Access	Sharing	Disposal
Confidentiality	Compromised sources or APIs exfiltrate feeds; weak TLS on collectors.	Misconfigured buckets or clusters; no encryption at rest; shared credentials.	Unencrypted staging tables; debug dumps with sensitive data.	Over-privileged BI roles; weak MFA; token theft.	Unvetted data-sharing, insecure transfer channels.	Unencrypted backups retained indefinitely; media disposed without wiping.
Integrity	Spoofed devices inject false events; log tampering.	Direct manipulation of stored datasets; unauthorised updates to reference data.	Data-poisoning of training pipelines; unaudited manual overrides.	SQL injection in ad-hoc tools; unauthorised edits via BI front-ends.	Partners altering shared datasets without detection; missing versioning.	Partial deletion leaving orphaned or inconsistent records.
Availability	DoS on ingestion endpoints; queue flooding.	Ransomware on storage clusters; loss of quorum.	Resource-exhaustion from hostile queries/jobs.	DoS against BI gateways; credential lockout abuse.	Dependency on fragile integration channels.	Destruction or corruption of backups or catalogs.
Privacy	Excess collection of personal data; absent consent.	Linkage across datasets enabling re-identification.	PII used in test environments; models memorising sensitive data.	Analysis can see personal data beyond remit; dashboards reveal small-cell aggregates.	Release of overly granular data; insufficient anonymisation or DP.	Retention beyond legal limits; failure to honour deletion requests.
AI-specific	Poisoned training streams; adversarial perturbations.	Injection of backdoored models into registries.	Model inversion/membership inference; compromised explainability tools.	Model-as-a-service endpoints abused for extraction; prompt-injection into BI agents.	Sharing models without protections enables cloning and misuse.	Legacy models not retired; logs of sensitive inputs retained indefinitely.
Insider	Ingestion operators bypass controls; manual uploads from personal devices.	DBAs or engineers misuse privileged access.	Data scientists export snapshots/models to personal storage.	Privileged users run broad exports; abuse of export features.	Deliberate over-sharing with favoured partners.	Insiders create off-site copies before deletion.

4. Zero Trust-based cybersecurity strategies for BI big data

4.1 Guiding Principles

Five principles shape the proposed strategies:

1. Data-centric security: controls follow data wherever it resides, not only at network perimeters. (Kantarcioglu & Ferrari, 2019)
2. Least privilege and continuous verification: access decisions depend on identity, context and data sensitivity, and are re-evaluated continuously. (Mushtaq, Mohsin, & Mushtaq, 2025)
3. Defence in depth: multiple, complementary controls reduce single points of failure. (Nair, 2021)
4. Secure-by-design analytics: security and privacy are built into pipelines, models and BI tools rather than bolted on. (Hussain & Hajjar, 2024)
5. Performance-aware design: cryptographic and monitoring choices recognise BI latency, throughput and cost constraints. (Taherdoost, Le, & Slimani, 2025)

4.2 Data Governance, Classification and Minimisation

Effective security begins with governance. Organisations should define:

- Data owners and stewards for each domain.
- Classification schemes (e.g., public, internal, confidential, highly restricted) linked to handling rules.
- Retention and disposal policies aligned with business and regulatory needs.
- Approval processes for new data sources and use cases, including privacy impact assessments. (Kantarcioglu & Ferrari, 2019)

Data minimisation—collecting only necessary data and retaining it only as long as required—is a central control for reducing exposure and complying with regulations such as GDPR. (Kantarcioglu & Ferrari, 2019) Governance mechanisms should therefore be embedded at ingestion (e.g., mandatory classification tags) and disposal (e.g., automated lifecycle policies and verifiable deletion).

4.3 Encryption and Key Management across the Data Lifecycle

Encryption underpins confidentiality and helps mitigate many cells in **Error! Reference source not found.** (especially storage and processing confidentiality/privacy threats):

- At rest: enforce strong encryption (e.g., AES-256) across object stores, file systems, databases and backups, with separate keys for different datasets or tenants. (Kantarcioglu & Ferrari, 2019)
- In transit: mandate TLS for all traffic between ingestion, storage, processing and BI components. Within service meshes, use mutual TLS (mTLS) to authenticate services and encrypt east-west traffic. (Mushtaq, Mohsin, & Mushtaq, 2025)
- In use: for highly sensitive workloads, consider homomorphic encryption or SMC to enable computation without exposing plaintext to counterparties, and TEEs to protect code and data during execution. (Taherdoost, Le, & Slimani, 2025), (Kantarcioglu & Ferrari, 2019)

Robust key management is essential: unique keys per dataset/tenant, automated rotation, strict access controls to key management services and full audit trails of key operations.

4.3.1 Design patterns for balancing cryptography and performance

Not all BI workloads need advanced cryptography. We propose three patterns:

1. Secure-by-default baseline

- Apply strong encryption at rest/in transit and robust IAM across all BI components.
- Suitable for most internal dashboards and operational analytics with tight latency budgets.

2. Selective advanced cryptography for high-risk workloads

- Use homomorphic encryption or SMC for cross-institution analytics where raw data cannot be shared (e.g., federated fraud detection across banks, multi-hospital outcome studies). (Taherdoost, Le, & Slimani, 2025)
- Restrict computations to aggregates where possible; complement with TEEs to improve performance while isolating sensitive code. (Kantarcioglu & Ferrari, 2019)

3. Privacy-preserving output for broad consumption

- Use differential privacy or risk-aware anonymisation primarily for public or widely shared BI outputs, where modest noise is acceptable. (Kantarcioglu & Ferrari, 2019)
- Keep internal operational dashboards minimally perturbed, relying on access control and monitoring to protect inputs.

These patterns allow organisations to reserve heavy cryptography for small but critical subsets of workloads, consistent with efficiency and profitability goals.

4.4 Identity, access management and fine-grained authorisation

To mitigate access- and insider-related cells in Table 3, BI environments need unified yet fine-grained IAM:

- Strong identities via a central identity provider with MFA for humans and certificate- or token-based identities for services.
- Least privilege enforced through a combination of RBAC and attribute-based policies tied to data classification, project, location, time and device posture. (Kantarcioglu & Ferrari, 2019)
- Row- and column-level security at the data layer, automatically applied to BI queries (e.g., restrict analysts to regional data; mask direct identifiers).
- Just-in-time (JIT) and just-enough access for privileged tasks, with automatic expiry and full logging.

By integrating IAM with classification metadata, access decisions become data-aware and context-sensitive—core Zero Trust properties.

4.5 Zero Trust BI reference architecture (ZT-BI)

Figure Y (conceptual) depicts the proposed ZT-BI architecture:

- Data sources: transactional systems, SaaS applications, IoT devices and external providers, connected via segmented source networks.
- Ingestion tier: batch and streaming collectors behind identity-aware gateways enforcing mTLS, rate limiting and schema validation.
- Storage tier: data lake/warehouse zones (raw, curated, sandbox, highly sensitive) in separate micro-segments, each with dedicated PEPs and encryption policies.
- Processing & analytics tier: ETL/ELT engines, streaming processors, ML platforms and notebook servers, connected through a service mesh providing mTLS, service identities and network policies.
- Access & presentation tier: BI portals, APIs and notebooks exposed through identity-aware proxies that perform MFA, device-posture checks and contextual risk scoring.
- Security & governance plane: PDPs, identity providers, key management, SIEM/SOAR, data catalogues and governance tools.

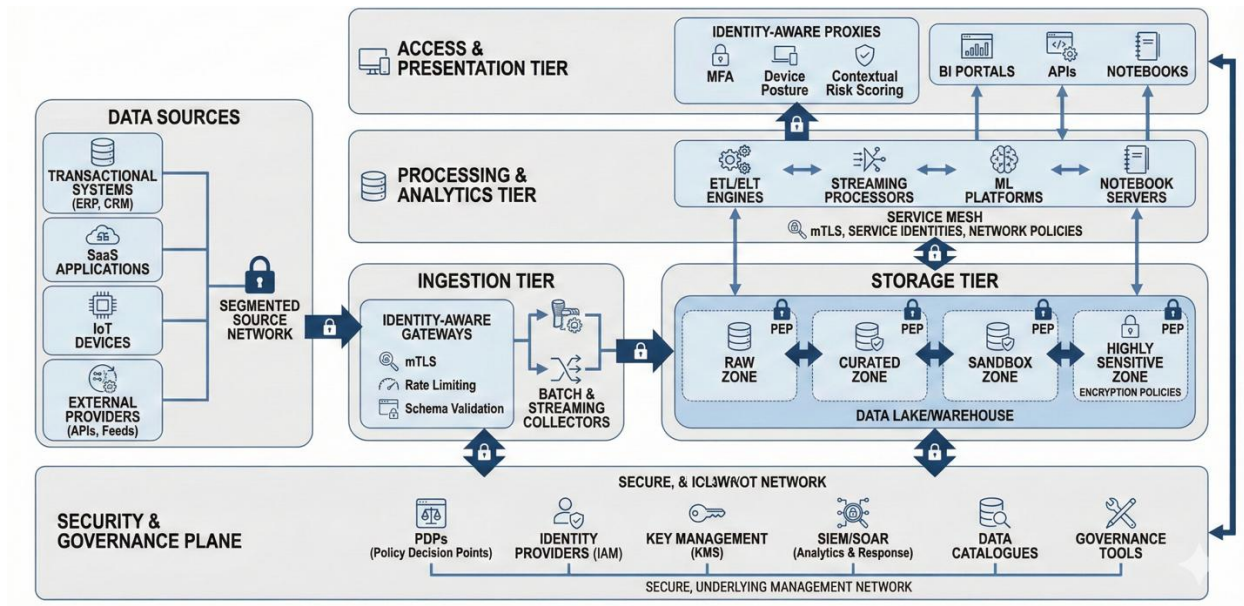


Figure 4 Illustration of the proposed Zero Trust Business Intelligence (ZT-BI) architecture and its integrated components.

This structure aligns with ZTA conceptual models such as the one in Mushtaq et al.'s Figure 2, which shows access requests flowing through PEPs and PDPs with continuous monitoring and context-aware trust evaluation. (Mushtaq, Mohsin, & Mushtaq, 2025)

4.6 Example scenario: sensitive financial query

Consider a data scientist building a fraud-detection model on highly sensitive transaction data:

1. The scientist logs into the BI platform; an identity-aware proxy enforces MFA and validates device posture. If risk indicators (new device, unusual country) are high, additional verification is required. (Nair, 2021)

2. On opening a notebook linked to the highly sensitive dataset, the PDP evaluates policies based on role, project, data classification and time. JIT access is granted for a limited session.
3. At query time, row-level filters restrict results to authorised regions and column-level masking hides direct identifiers.
4. Aggregates destined for a broader audience are post-processed with differential privacy before being written to a curated analytics table. (Kantarcioglu & Ferrari, 2019)
5. All queries and data movements are logged. An AI-driven anomaly detector watches for unusual access patterns (e.g., large exports, unusual joins). If detected, alerts are raised and sessions can be throttled or terminated. (Hussain & Hajjar, 2024)

This scenario illustrates how principles and components of ZT-BI jointly mitigate confidentiality, privacy, AI-specific and insider threats across processing and access stages.

4.7 AI-driven security analytics

Given the scale and complexity of BI platforms, AI-based security analytics are needed to detect patterns humans would miss:

- Anomaly detection on query patterns, login behaviour, data transfers and resource usage. (Hussain & Hajjar, 2024)
- Threat-intelligence integration correlating internal telemetry with known indicators of compromise.
- Automated response via SOAR tools that revoke tokens, disable accounts or isolate workloads based on detection outputs. (Mushtaq, Mohsin, & Mushtaq, 2025)

These models must themselves be robust against poisoning and evasion, and—where they influence critical access decisions—should employ explainable AI techniques for transparency and auditability. (Taherdoost, Le, & Slimani, 2025)

4.8 Privacy-preserving analytics and regulatory compliance

Privacy-preserving techniques complement security controls:

- Pseudonymisation/tokenisation of direct identifiers with secured mapping tables.
- Differential privacy and risk-aware anonymisation for shared or published datasets. (Kantarcioglu & Ferrari, 2019)
- Data subject rights support: ability to locate, export, rectify and delete an individual's data across stores—critical under GDPR and similar laws.

These measures primarily mitigate privacy and AI-specific cells in Table 4 *Example maturity characteristics by domain*, particularly in sharing and disposal stages.

4.9 Integrity assurance and tamper-evident logging

To protect integrity and support non-repudiation:

- Maintain append-only, cryptographically linked logs of data lineage, transformations and access events.
- For high-value multi-party use cases, consider blockchain or distributed ledgers for shared integrity proofs and smart-contract-based access rules. (Hussain & Hajjar, 2024)

Given overhead and complexity, blockchain is best reserved for scenarios where multi-party auditability and trust minimisation justify the cost.

5. Implications for Operational Efficiency and Profitability

Strong security is sometimes seen as a drag on BI agility. However, empirical evidence suggests the opposite when controls are well designed. Egerson et al. report that cybersecurity capabilities significantly predict operational efficiency and profitability in Nigerian deposit money banks. (Egerson, Williams, Aribigbola, Okafor, & Olaleye, 2024) Mechanisms include:

- Reduced downtime: resilience against ransomware and DoS protects digital channels and BI tools.
- Higher decision quality: integrity controls reduce the risk of decisions based on corrupted data.
- Customer trust and regulatory compliance: sound security and privacy practices protect reputation and avoid fines.
- Operational insights from security analytics: monitoring can reveal misconfigurations and inefficiencies beyond security issues.

In Williams' broader review of AI adoption, responsible security, privacy and governance are presented as enablers of sustainable AI-driven efficiency gains rather than obstacles. (Williams, 2024)

The design patterns in section 4 further support efficiency by limiting heavy cryptography to high-risk, latency-tolerant workloads and using lighter but robust controls elsewhere.

6. Implementation Roadmap and Zero Trust Maturity for BI

Building on ZeTuMM, (Modderkolk, 2018) we propose a BI-specific maturity model with four levels across five domains: Permission, Infrastructure, Processes, Intelligence and People.






DOMAINS	MATURITY LEVELS			
	LEVEL 1 – INITIAL: Fragmented controls, ad hoc processes, little linkage to data lifecycle.	LEVEL 2 – MANAGED: Consistent basic protections across major BI components.	LEVEL 3 – DEFINED: Organisation-wide Zero Trust-aligned policies and architectures covering the BI stack.	LEVEL 4 – ADAPTIVE: Controls dynamically tuned by telemetry and risk; continuous improvement integrated with business KPIs.
 PERMISSION	Fragmented controls	Basic protections	Aligned policies	Dynamic tuning
 INFRASTRUCTURE	Flat networks, inconsistent encryption	Encryption at rest/in transit	Micro-segments, PEPs, Service meshes, mTLS	Hardened configurations, secure baselines
 PROCESSES	Ad hoc incident response, limited governance	Limited governance	Formal data governance, documented IR, regular exercises, clear data lifecycle	Continuous improvement, integrated with business KPIs
 INTELLIGENCE	Sparse logging, manual monitoring	Centralized logging	AI-driven anomaly detection, external threat intelligence, feedback loops	Dynamic tuning by telemetry/risk
 PEOPLE	Ad hoc training	Basic awareness	Secure coding training, role-specific awareness, performance metrics	Continuous improvement, rewarding secure behavior

Figure 5 A four-level BI Zero Trust Maturity Model showing progression across five key domains.

6.1 Maturity levels

To operationalise Zero Trust principles within business intelligence ecosystems, organisations require a structured mechanism for assessing their current posture and guiding progressive enhancement. A maturity model provides this structure by articulating clearly defined stages of capability development, allowing organisations to benchmark their practices and set realistic improvement trajectories. Building on the conceptual foundations of existing Zero Trust maturity models and adapting them to the unique characteristics of big-data-driven BI environments, four maturity levels can be distinguished. These levels reflect increasing coherence, automation, contextual awareness and strategic alignment across governance, technology and human processes, culminating in a fully adaptive BI security ecosystem that continuously evolves in response to changing risks.

- Level 1 – Initial: fragmented controls, ad hoc processes, little linkage to data lifecycle.
- Level 2 – Managed: consistent basic protections across major BI components.
- Level 3 – Defined: organisation-wide Zero Trust-aligned policies and architectures covering the BI stack.
- Level 4 – Adaptive: controls dynamically tuned by telemetry and risk; continuous improvement integrated with business KPIs. (Mushtaq, Mohsin, & Mushtaq, 2025)

6.2 Maturity characteristics by domain

Table 4 Example maturity characteristics by domain

Domain	Level 1 – Initial	Level 2 – Managed	Level 3 – Defined	Level 4 – Adaptive
Permission (IAM & authorisation)	Identity silos; shared accounts; coarse DB-level grants.	Central IdP with MFA; basic RBAC for key datasets.	Unified identities for humans/services; ABAC tied to data classification; row/column-level security in BI tools; JIT elevation.	Risk-based authentication; continuous right-sizing of roles; automated detection/remediation of excessive privileges.
Infrastructure (network, storage, compute)	Flat networks; inconsistent encryption; manual .configuration	Segmented environments for major tiers; encryption at rest .in main stores	Micro-segmented ZT-BI architecture; mTLS across east–west traffic; hardened baselines and config management.	Policy-driven infrastructure-as-code; automated micro-segmentation based on data flows; continuous posture assessment and self-healing.
Processes (governance & IR)	Informal or absent processes; unclear data .ownership	Named data owners; basic incident runbooks for key BI .platforms	Formal data governance board; lifecycle policies; integrated incident response covering pipelines and models.	Threat-hunting focused on BI assets; regular red-teaming; governance and IR metrics feed continuous improvement.
Intelligence (monitoring & analytics)	Limited logging; manual review; no BI-specific .coverage	Central log collection; rule-based alerts for BI .components	AI-based anomaly detection on BI queries, access and data flows; SOAR playbooks for common scenarios.	Behavioural models retrained continuously; feedback from incidents and business outcomes refines detection and policies.
People (skills & culture)	Minimal security awareness; BI staff unaware of .obligations	Generic awareness training; occasional BI-specific .sessions	Role-specific training for data engineers, scientists and analysts; security responsibilities in job descriptions.	Security-by-design culture; cross-functional “purple teams” including BI; incentives for reducing risk and improving secure practices.

6.3 Example metrics

To make the model measurement-ready, organisations can track:

Permission: percentage of users with standing admin privileges; proportion of privileged sessions using JIT elevation; number of orphaned service accounts.

Infrastructure: percentage of data stores with encryption at rest; proportion of east–west BI traffic protected by mTLS; mean time to patch critical vulnerabilities on analytics clusters.

Processes: number and severity of BI-related incidents per quarter; proportion of new BI projects that undergo privacy impact assessment.

Intelligence: mean time to detect and respond to BI-related incidents; percentage of BI components emitting logs to the central SIEM; coverage of anomaly-detection models across BI workloads.

People: proportion of BI staff completing role-specific security training; survey-based indicators of security culture.

These metrics give quantitative complements to the qualitative descriptions in Table 4 and can be tied to business KPIs such as revenue per channel, incident-related downtime and regulatory findings.

7. Limitations and challenges

Zero Trust migration in BI environments faces several constraints:

- Technical debt and legacy systems. Many organisations operate legacy warehouses, ETL tools and bespoke reporting systems that are difficult to retrofit with fine-grained access control, micro-segmentation and pervasive encryption without substantial re-engineering.
- Skills scarcity. Expertise in ZTA design, advanced cryptography and AI security engineering is limited; misconfigurations can undermine intended protections.
- Cost and complexity of advanced mechanisms. Blockchain-based integrity solutions and privacy-preserving multi-party analytics introduce overhead and may be disproportionate in many settings.
- Organisational resistance. Changes in access patterns, logging and governance can encounter cultural resistance, particularly when they constrain long-standing practices.

Methodologically, this paper is based on an integrative review and architectural synthesis. While it draws on empirical evidence of cybersecurity's impact on efficiency and profitability, it does not provide new quantitative evaluation of the proposed framework. Future empirical case studies are needed to measure the actual effects of ZT-BI implementations in different sectors.

8. Limitations and challenges

Big-data-driven BI systems sit at the core of organisational decision-making and innovation, yet they also concentrate sensitive information and critical infrastructure in ways that attract sophisticated adversaries and raise significant privacy and national-security concerns. This paper has argued that protecting big data in BI environments requires moving beyond perimeter-based security models toward a Zero Trust, data-centric paradigm. In doing so, it introduced a BI-specific threat model and lifecycle matrix that systematically links major categories of threats to each stage of the data lifecycle, and it presented a Zero Trust BI reference architecture (ZT-BI) supported by strategies in governance, encryption, identity and access management, micro-segmentation, AI-driven monitoring, privacy-preserving analytics and integrity assurance. The paper also developed a maturity roadmap and accompanying metrics tailored specifically to BI ecosystems, extending and operationalising concepts from ZeTuMM for real-world implementation.

The framework demonstrates that well-designed cybersecurity capabilities can enhance, rather than hinder, operational efficiency and profitability by reducing downtime, improving the reliability of decision-making and strengthening regulatory compliance. Looking ahead, future research should empirically evaluate ZT-BI deployments to understand their measurable effects on security incidents, BI performance and broader business outcomes. Additional investigation is needed into practical tooling that can integrate advanced cryptographic techniques—such as homomorphic encryption, secure multiparty computation, trusted execution environments and differential privacy—into mainstream BI platforms without imposing prohibitive complexity or performance costs. Research should also focus on developing robust and explainable AI methods for security analytics that remain resilient in the face of adversarial manipulation, as well as refining regulatory and ethical frameworks that can balance innovation, privacy protection, economic value and national-security imperatives in big-data-driven BI contexts.

As organisations continue to deepen their reliance on BI and AI, the alignment of cybersecurity, data governance and analytics strategy will become increasingly critical. A Zero Trust-based approach, implemented with careful attention to both security and performance, offers a sustainable path for leveraging big data responsibly, effectively and safely.

9. References

- Egerson, J. I., Williams, M., Aribigbola, A., Okafor, M., & Olaleye, A. (2024). Cybersecurity strategies for protecting big data in business intelligence systems: Implication for operational efficiency and profitability. *World Journal of Advanced Research and Reviews*.
- Hussain, D., & Hajjar, L. (2024). Cybersecurity and Big Data Analytics: Strategies for Securing Business Intelligence in the Digital Era. Research Gate.
- Kantarcioglu, M., & Ferrari, E. (2019). Research Challenges at the Intersection of Big Data, Security and Privacy. Texas: Frontiers in Big Data.
- Modderkolk, M. (2018). *Zero Trust Maturity Matters: Modeling Cyber Security Focus Areas and Maturity Levels in the Zero Trust Principle* . Utrecht: Department of Information and Computer Science.
- Mushtaq, S., Mohsin, M., & Mushtaq, M. M. (2025). A Systematic Literature Review on the Implementation and Challenges of Zero Trust Architecture Across Domains. *Sensors*. MDPI.
- Nair, A. (2021). The Why and How of adopting Zero Trust Model in Organizations. Reserch Gate.
- Stephan, P. B. (2025). Big Data as a National Security Issue. University of Chicago.
- Taherdoost, H., Le, T.-V., & Slimani, K. (2025). Cryptographic Techniques in Artificial Intelligence Security: A Bibliometric Review. *Cryptography*. MDPI.
- Williams, M. T. (2024). Future visions of ai enhancing industries and navigating ethical landscapes. *International Journal of Science and Research Archive*.