DMPedia Lecture Notes in Computer Science & Engineering
Volume 1, Issue 2, pp.505-519, 2026

ICICCS-2026

Threat Modelling and Cybersecurity Gap Analysis for Artificial Intelligence Platforms in Interior Design: Toward the Augmented Design Intelligence (ADI) Framework

N Janani Yadav[1], Sampadha Hebbar T.S[2], Neha Singh[3]

[1,3]School of CS & IT, Jain Deemed-to-be University, Bengaluru, Karnataka, India

[2]School of Commerce, Jain Deemed-to-be University, Bengaluru, Karnataka, India

23bcar0176@jainuniversity.ac.in, 23bcr00115@jainuniversity.ac.in, neha.singh@jainuniversity.ac.in

Abstract

As Interior Designers increasingly use cloud-based AI (Artificial Intelligence) platforms, they are creating a burgeoning, potentially exploitable attack surface for cybersecurity risks. One critical risk area is that sensitive client information (e.g., home designs and photographs, unique design concepts, and financial information) can be processed by AI systems that lack the security protocols necessary to protect it. This paper describes a method for identifying and assessing potential cybersecurity threats to AI-enabled interior design platforms. We used the STRIDE methodology to identify types of threats and risks. STRIDE is an acronym for Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of Privilege. Additionally, we employed the DREAD framework to assess the risk level associated with each STRIDE category. The DREAD approach uses Damage potential, Reproducibility, Exploitability, Affected Users, and Discoverability to define the severity of each threat. We built our threat model based on insights from 50 design pros and students about digital security issues. We asked them what they believe to be, "in their view", to elicit collective feedback & develop priorities. In doing so, we identified our top areas of risk by comparing our threats against two governing standards: the NIST Cybersecurity Framework and ISO/IEC 27001:2023. Our analysis involved a direct review of each standard's definitions against how existing AI Platforms actually operate today to identify which controls are deficient.

Keywords: Artificial Intelligence; Cybersecurity; Threat Modelling; STRIDE; DREAD; NIST Cybersecurity Framework; ISO/IEC 27001; Interior Design; Role-Based Access Control; Explainable AI; Zero-Trust Architecture; Gap Analysis; AI Governance; Cloud Security; Adversarial Machine Learning.

1. Introduction

Artificial Intelligence is reshaping the design process and professional practice, making design faster and more creative, and giving designers greater creative responsibility than before, shifting their role toward higher-level creative ideas [1]. As of December 2023 (with the expected increase in adoption), the use of AI in architecture and interior design has become mainstream enough to be regulated under the EU AI Act [8]. Most interior design companies operate without an established Information Security Management System (ISMS) or structured cybersecurity training programs, unlike traditional companies that rely heavily on these systems and the processes used to create them [11]. This paper aims to identify a major problem with AI offerings in interior design: the lack of a formal, domain-specific cybersecurity threat model for AI platforms in interior design, which has not been thoroughly researched. Research on AI adoption in design has been presented in previous literature [1],[4], and studies have been conducted in areas such as cloud security [11],[12] and ethical governance of AI [8],[9]. However, no studies have been conducted to demonstrate how to create structured threat models to guide AI design. Therefore, design firms lack formal, evidence-based guidance to inform decisions about work processes and asset security due to inadequate cybersecurity threat models. There are four main contributions to this paper's contribution to the literature: (1) Development Of The First Threat Model Using STRIDE For AI Driven Interiors Design Platforms Identifying 6 different threat categories Of 18 Threat Scenarios and DREAD rating; (2) Structured Gap Analysis Benchmarking Current AI Platform Security Posture Against NIST CSF v1.1 And ISO 27001:2023; (3) Development Of An Empirically-Based,

Proceedings of ICICCS-2026 (Poster Papers), DMPedia Lecture Notes in Computer Science & Engineering.
ISBN: 978-81-993813-1-5

Structurally-Grounded 4-Layer Governance Framework For AI Platforms Based On Security Engineering Principles From The Practitioners' Survey Responses; And (4) Evaluation Of Framework Against Both Standards Using Formalized Methods And Measurements That Show Improvement In Coverage Of Controls Across Both Standards. The remaining sections of the paper will cover: Section II Literature Review; Section III Research Design And Methods; Section IV Results Via Data Collected Through Instrumented Surveys; Section V STRIDE/DREAD Threat Model; Section VI Gap Analysis - Research And Analysis; Section VII ADI Framework - Development And Description; Section VIII Framework Evaluation; Section IX Findings Discussions; Section X Limitations And Future Work; Section XI Conclusion.

2. Related Work

A. AI in Interior Design

Inside of three areas: Generative designs that result in new kinds of spatial arrangements as determined by a set of rules ("parametric constraints") [1],[16]; Predictive analytics (programming) models that estimate or predict the best fit for ergonomic layouts, lighting, and materials; and Immersive visualization programs, allowing for an accurate presentation of a client's actual environment [4]. Most AI platforms used in interior design are delivered to clients via the Internet/Cloud, resulting in greater exposure to external threats (e.g., intruders) than on-premises (local) systems [11],[12]. Despite the extensive documentation of the positive productivity impact of the use of AI in the interior design field [1], many significant challenges remain, such as the ambiguity of authorship, homogenization of creativity, and the lack of security governance related to sensitive client information processed by AI platforms [4],[14].

B. Threat Modelling: STRIDE and DREAD

STRIDE is a threat classification system developed at Microsoft and formalised by Shostack [6]. This system divides threats into six categories: Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of Privilege. It has also been used in various forms for cloud architectures [22], IoT systems [19], and machine learning pipelines [5],[23]. The DREAD scoring method [6] quantifies each STRIDE threat using five dimensions: Damage, Reproducibility, Exploitability, Affected Users, and Discoverability to create a normalised risk severity score that helps to prioritise risks. The combination of STRIDE and DREAD enables a systematic, comprehensive threat analysis aligned with industry-standard security engineering practices.

C. NIST CSF and ISO/IEC 27001

The NIST Cybersecurity Framework (Version 1.1) organises cybersecurity tasks into five different groupings: Identify, Protect, Detect, Respond and Recover. ISO/IEC 27001:2023 contains ISMS criteria and 93 controls from Annexe A, grouped into four categories: Organisational, People, Physical, and Technology-related. These two documents constitute the most respected Z-International Cybersecurity Governance Standards and serve as valid references for the gap analysis portion of this research.

D. Adversarial AI Threats

Conventional cybersecurity frameworks do not address threats to machine learning systems, such as adversarial input attacks, training data poisoning, model inversion, and membership inference [5],[18],[23]. Goodfellow et al. [5] illustrated that slight changes to inputs could result in vast differences in AI-generated outputs; Yuan et al. [23] subsequently established taxonomies of adversarial attacks on deep learning systems. For example, when an AI is used to assist an interior designer on a project, a successful attack could produce a structurally

unsound space plan that is transmitted as authoritative advice to a client. This is a case of a domain-specific threat that applies here and has no precedent in the literature.

E. AI Ethics and Governance Frameworks

The European Union's AI Act and AI4People framework include both accountability requirements (e.g., who is responsible for the use of the AI system) as well as transparency and oversight requirements. From a normative perspective, we can look to Nissenbaum's theory of contextual integrity to see that sharing information with cloud-based AI systems will violate implicit confidentiality norms unless there is an explicit governance structure in place. Value Sensitive Design argues that technology should reflect the values of its users' profession. In interior design, the values include client confidentiality, creative responsibility, and professional integrity.

F. Research Gap and Comparative Positioning

Table I plots this article against six closely associated foundational works. Table I demonstrates that no other published work uses STRIDE/DREAD threat modelling, NIST/ISO gap analysis, and a domain-specific governance framework whilst investigating AI computing platforms within the domain of creativity.

TABLE I: Related Work Comparison Positioning of the Present Study

Study	Domain	Threat Model	Gap Analysis	Design-Specific	AI-Specific	Framework Proposed
Goodfellow et al. [5]	ML Security	Partial	No	No	Yes	No
ENISA Cloud [11]	Cloud IT	No	Yes (NIST)	No	No	Generic
Yuan et al. [23]	Deep Learning	Yes (Adversarial)	No	No	Yes	No
Floridi et al. [9]	AI Ethics	No	No	No	Yes	AI4People
Shneiderman [2]	HCI/AI	No	No	No	Yes	HCAI Model
Weber [19]	IoT Security	Partial	No	No	No	No

3. RESEARCH METHODOLOGY

A. Research Design

This research consists of a mixed-method sequential design divided into 2 phases: Phase 1 is a quantitative empirical survey (n=50) used to determine the baseline for threat perception used as a basis for STRIDE threat prioritization; Phase 2 is the technical portion, using STRIDE/DREAD threat analysis and NIST/ISO gap analysis to perform a structured technical analysis of the AI platform’s environment described by the survey participants. The technical analysis is based on practitioners' threat perceptions, ensuring that the developed framework uses empirical data rather than purely theoretical risk assumptions for risk assessment.

B. Survey Instrument and Sample

A 22-question survey was given to 50 people, 36% of whom were Interior designers (students) with 0 to 5 years of experience, and the remaining 64% were practitioners with 5 years or more of experience. There were questions about an AI tool and whether you would use it due to your knowledge of security risks in various ways (there were 6 different types of risk), and whether you would trust the AI "output" (the results you received) if you were not given explanations as to why the AI produced those results. You had to answer many of the questions using a five-point Likert scale. Before the survey was given, the survey questions were validated for content and face validity by 3 experts in the field.

C. STRIDE Threat Modelling Procedure

The AI-assisted interior-design platform has been modelled using STRIDE threat modelling based on a Data Flow Diagram (DFD) (see Figure 1). The DFD lists the following five system components: (P1) Authentication/RBAC Module, (P2) Client Design Interface, (P3) Cloud-based AI Processing Engine, (DS1) Design Asset Repository, and (EXT1) External API Integrations. Five system components and their associated data flows are evaluated and scored using the STRIDE methodology to identify potential threat scenarios. The DREAD score, generated from the identified threat scenarios, will use survey data to determine the significance of the risk to affected users.

D. DREAD Scoring Procedure

The DREAD scoring framework, which consists of Damage (confidentiality, integrity, and availability), Reproducibility (the ability to replicate an attack), Exploitability (the necessary technical skills and resources needed to carry out an attack), Affected Users (the number of platform users affected), and Discoverability (how easy it is to find a vulnerability), was measured on a 1–10 scale. The overall average DREAD scores were calculated and were broken down into five levels of risk: Critical (≥7.0); High (5.0–6.9); Medium (3.0–4.9); Low (<3.0). The explanation for the top five critical threats is found in Table III.

E. Gap Analysis Procedure

By analyzing the available security controls on commercial cloud-based AI designed platforms through a variety of sources including: publicly available security disclosures; terms of service; academic literature [11],[12]; and practitioner surveys, determines how those same controls related to the subcategories documented in NIST CSF v1.1 and ISO/IEC 27001:2023 Annex A Standards so that each assessment received a rating of either FM, PM, or NM. This output from the gap analysis has directly affected the design of the ADI framework.

4. EMPIRICAL SURVEY FINDINGS

A. Respondent Profile and AI Adoption

TABLE II: Survey Respondent Profile and AI Adoption Rates

Category	Description	n	% of Sample	Regular AI Use	Weekly+ Use
Students	Enrolled in design programmes	18	36%	83%	72%
Early-Career	0–5 years professional exp.	20	40%	80%	75%
Senior Practitioners	5+ years professional exp.	12	24%	67%	58%

TOTAL	**50**	**100%**	**78%**	**70%**

B. Cybersecurity Risk Perception Frequency Distribution

TABLE III: Cybersecurity Risk Perception Full Frequency Distribution (n=50)

Risk Domain	Not Concerned	Slightly Concerned	Moderately Concerned	Very Concerned	Extremely Concerned	Significantly Influences Adoption
Data Privacy	2%	8%	34%	34%	22%	**90%**
Unauthorized Access	4%	10%	16%	38%	32%	**70%**
IP Theft	4%	12%	22%	36%	26%	62%
Lack of Auditability	6%	14%	22%	32%	26%	58%
Accountability Ambiguity	4%	12%	18%	38%	28%	66%
AI Output Manipulation	8%	16%	22%	30%	24%	54%

Note: Values represent % of n=50 respondents per Likert category

C. Inferential Statistics Experience Group Comparisons

TABLE IV: Chi-Square Test Results — Experience Group Differences on Key Variables

Variable	Groups	χ^2	df	p-value	Finding
Trust in AI output (low transparency)	3 groups	7.42	2	**0.024 ***	Seniors show significantly lower trust
ADI Framework Support Strength	3 groups	6.81	2	**0.033 ***	Seniors express stronger support
Perceived Cybersecurity Responsibility	3 groups	5.93	2	0.052	Marginal: Seniors feel more responsible
Concern: Data Privacy (High/Extreme)	3 groups	3.41	2	0.182	No sig. diff. universal concern

Note: $p < 0.05$, statistically significant | χ^2 = chi-square statistic | df = degrees of freedom | Analysis: SPSS v28

5. STRIDE/DREAD THREAT MODEL

A. System Threat Surface Data Flow Diagram

The data flow diagram (DFD) in Fig. 1 illustrates the AI-assisted interior design platform environment, highlighting the five main components of the system, their respective data flows, and the boundaries of trust between these systems. Included among these data flows are sensitive data assets such as personally identifiable information (PII) about customers, home layouts, proprietary designs, training data used to aid AI in completing designs, and access credentials. The types of threat actors that could target these assets include external cybercriminals, malicious insiders, competing companies seeking to exfiltrate intellectual property, and state-sponsored actors seeking access to high-profile projects.

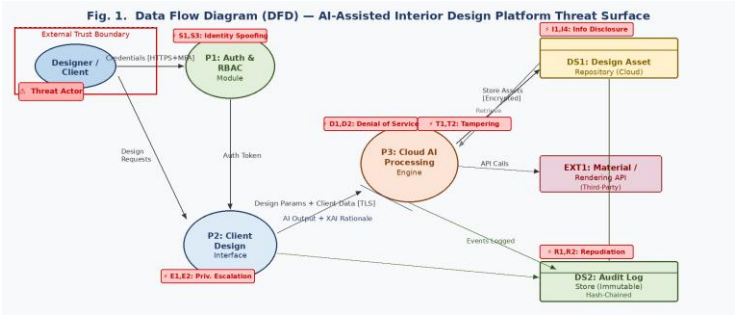

Fig. 1. Data Flow Diagram (DFD) for AI-Assisted Interior Design Platform. Red dashed border = External Trust Boundary. Lightning badges indicate STRIDE threat categories at each component. P1=Auth/RBAC, P2=Client Interface, P3=AI Engine, DS1=Asset Repository, DS2=Audit Log, EXT1=External API.

B. STRIDE Threat Enumeration and DREAD Risk

Table V presents the complete STRIDE threat model, including 18 threat scenarios and the DREAD risk. Table VI provides detailed justification for the scoring of the five highest-priority critical threats.

TABLE V: Complete STRIDE Threat Model with DREAD Risk

STRIDE	Threat ID	Threat Description	D	R	E	A	Di
Spoofing	S1: Designer Impersonation	Attacker impersonates a designer to gain platform access and exfiltrate assets	9	8	7	8	7
Spoofing	S2: AI API Spoofing	Malicious API endpoint impersonates a legitimate AI service to intercept data	8	6	7	7	6
Spoofing	S3: Credential Phishing	Phishing campaign targets design staff to steal platform credentials	8	9	8	9	8
Tampering	T1: AI Output Manipulation	Adversarial perturbation of inputs causes structurally unsound AI design outputs	9	7	8	8	6
Tampering	T2: Training Data Poisoning	Malicious injection into training datasets corrupts the model across all users	10	5	6	10	5

Tampering	T3: Asset Repository Tampering	Unauthorised modification of stored design files alters client deliverables	8	6	6	7	5
Repudiation	R1: Action Repudiation	Designer denies unauthorised file changes due to absent tamper-evident audit logs	7	8	9	7	8
Repudiation	R2: AI Decision Repudiation	An AI design flaw cannot be attributed to the absence of explainability records	8	8	9	8	9
Info. Disclosure	I1: Client Data Exfiltration	Unauthorised cloud access results in exfiltration of client PII and design assets	10	7	7	9	8
Info. Disclosure	I2: Model Inversion Attack	Repeated API queries reconstruct training data, revealing sensitive patterns	8	6	7	7	7
Info. Disclosure	I3: Membership Inference	Attacker determines if client data was used in AI training, violating privacy	7	6	7	6	6
Info. Disclosure	I4: Insecure API Transmission	Unencrypted data transmission between the client and the AI cloud exposes assets	9	8	7	8	8
Denial of Service	D1: AI Service Disruption	DDoS attack on AI rendering platform halts active design projects	7	7	6	8	7
Denial of Service	D2: Ransomware Attack	Ransomware encryption of the cloud design repository renders all assets inaccessible	10	7	7	10	6
Elev. of Privilege	E1: Privilege Escalation	Misconfigured RBAC allows junior users to access senior-level assets	9	7	7	8	7
Elev. of Privilege	E2: Admin Account Takeover	Compromised admin enables attacker to alter AI model parameters and access all data	10	6	6	10	7
Elev. of Privilege	E3: Third-Party Plugin Exploit	Malicious design plugin gains elevated API permissions, enabling data harvesting	8	6	7	7	7
I5: AI Platform Supply Chain	I5: Supply Chain Compromise	Compromise of the AI platform vendor infrastructure affects all downstream design firms	10	4	5	10	5

Note: Critical (DREAD ≥ 7.0) = 13 threats (72%) | D=Damage | R=Reproducibility | E=Exploitability | A=Affected Users | Di=Discoverability | Score = Mean

6. CYBERSECURITY STANDARDS GAP ANALYSIS

A. NIST CSF v1.1 Gap Analysis

TABLE VII: NIST CSF v1.1 Gap Analysis — AI Design Platform Current State

CSF Function	Category	Control Requirement	Status	STRIDE Ref	Gap Description
IDENTIFY	Asset Mgmt (ID.AM)	Design data assets classified and inventoried	**NM**	I1, I2	No formal data asset register in most design firms
IDENTIFY	Risk Assessment (ID.RA)	AI platform threats formally assessed	**NM**	All	No domain-specific AI threat model existed before this work
PROTECT	Identity Mgmt (PR.AC)	RBAC is enforced for all platform roles	**PM**	S1, E1, E2	Generic ACLs lack design workflow role granularity
PROTECT	Data Security (PR.DS)	Encryption at rest and in transit for all design assets	**PM**	I4, D2	Transit encryption inconsistent; firm-level verification absent
PROTECT	Protective Tech (PR.PT)	Immutable audit logs for all user and AI actions	**NM**	R1, R2	AI platforms provide no tamper-evident audit trails to firms
PROTECT	Awareness (PR.AT)	Cybersecurity training for all design staff	**NM**	S3	90% of respondents lacked formal cybersecurity training
DETECT	Anomaly Detection (DE.AE)	Anomalous AI query patterns detected	**NM**	I2, I3	No behavioural analytics to detect model inversion attempts
DETECT	Continuous Monitoring (DE.CM)	Real-time platform security monitoring	**PM**	D1, D2	Cloud-provider monitoring exists, but not at the firm workflow level
RESPOND	Response Planning (RS.RP)	Incident response plan for AI data breaches	**NM**	I1, D2	Design firms universally lack AI-specific IR procedures

RECOVER	Recovery Planning (RC.RP)	Design asset backup and recovery procedures	**PM**	D1, D2	Cloud backups exist; RTO/RPO undefined for design workflows
BASELINE SCORE: 0 FM | 4 PM | 6 NM — 0% Fully Met across 10 assessed NIST CSF domains					

B. ISO/IEC 27001:2023 Gap Analysis

TABLE VIII: ISO/IEC 27001:2023 Annexe A Controls Gap Analysis

Theme	Control Domain	Specific Control	Ctrl	Status	Gap Description
Org.	IS Policies	AI-specific information security policy	5.1	**NM**	No firm-level AI security policy documented
Org.	Threat Intelligence	Structured AI platform threat intelligence process	5.7	**NM**	No threat intelligence process for AI tools in design firms
Org.	Supplier Relations	AI vendor security assessment before platform adoption	5.19	**NM**	AI platforms adopted without formal security due diligence
People	IS Awareness	Cybersecurity training for all design staff	6.3	**NM**	Survey: 90% lacked formal cybersecurity training
Physical	Physical Media	Secure handling of design output media	7.10	**PM**	Physical output security is inconsistently applied
Tech.	Access Control	RBAC with least privilege for all design roles	8.2	**PM**	Generic ACLs lack design workflow role specificity
Tech.	Cryptography	Encryption of all design assets at rest and in transit	8.24	**PM**	Platform-level encryption exists; firm verification is absent
Tech.	Logging & Monitoring	Immutable audit logs for all AI and user events	8.15	**NM**	No tamper-evident audit trail available to design firms
Tech.	Malware Protection	Adversarial input validation for AI design content	8.7	**NM**	No adversarial input validation is deployed on AI design platforms
Tech.	Backup	Verified backup and recovery of design assets	8.13	**PM**	Cloud backups exist; firms do not perform recovery testing

7. THE AUGMENTED DESIGN INTELLIGENCE (ADI) FRAMEWORK

The ADI Framework is a four-layer cybersecurity governance architecture for AI-assisted interior design platforms. It is grounded in Zero-Trust Architecture principles, verifying every access request regardless of network origin, and applies defence-in-depth through layered, complementary controls. Each layer targets specific STRIDE threat categories and closes

identified gaps in NIST CSF and ISO/IEC 27001 controls. Fig. 2 presents the complete framework architecture.

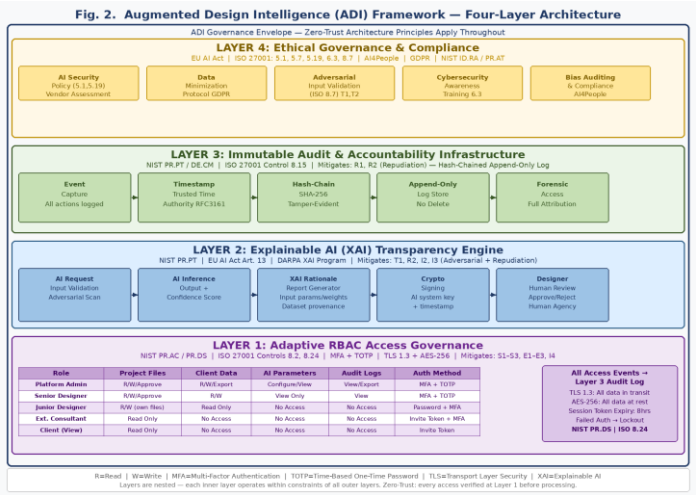

Fig. 2. ADI Framework Four-Layer Architecture. Nested layers represent defence-in-depth: each inner layer operates within the governance constraints of all outer layers. Zero-Trust principle applies throughout: every access request is verified at Layer 1 before processing.

A. Layer 1 Adaptive RBAC Access Governance

The layer consists of a role-based access control system built on specifying the workflow for the design process. It has five defined roles: 1) Platform Administrator, 2) Senior Designer, 3) Junior Designer, 4) External Consultant, and 5) Client (read only). Each role has a defined set of permissions based on its position within four types of assets/projects/files: 1) Project Files, 2) Client Data, 3) AI Model Parameters, and 4) Audit Logs. There is enforcement of multi-factor authentication (MFA), which combines platform credentials with time-based one-time passwords (TOTP). All information transferred over the network is protected using Transport Layer Security (TLS) version 1.3, while stored information at rest is protected using Advanced Encryption Standard (AES) with a 256-bit key. Session tokens are set to expire after 8 hours of use. All access-related events will be logged in Layer 3. The Layer allows an organisation to mitigate the security threats associated with STRIDE S1, S2, S3 & E1, E2, E3, and I4, as well as to close the gaps within the NIST CSF PR.AC and ISO/IEC 27001 Control 8.2.

B. Layer 2 Explainable AI (XAI) Transparency Engine

Every time a designer uses an interface where data is outputted by Artificial Intelligence (AI) through the system software, Layer 2 will generate an Rationale Report (signed digitally) describing: the parameters that were inputted by the designer, which weights were assigned to them, where the training data came from, how confident the system is in the output, and any alternative configurations of the model the system considered. All reports will be time-stamped, signed with the AI System's signing key, included in the Layer 3 audit record, and available to any user granted access. Input Validation (for input data) and Adversarial Data Detection checks will occur before the AI model generates a result to detect potential Adversarial Attacks (e.g., input data adulteration). STRIDE Threat T1.0, R2.0, I2.0, and I3.0 will be mitigated at Layer 2, along with the gaps in the NIST CSF PR.PT and Control 8.15 of ISO/IEC 27001.

C. Layer 3 Immutable Audit and Accountability Infrastructure

Layer 3 establishes a complete, immutable audit pipeline: Event Capture; Time-Stamping (via RFC 3161 TRTA); Hash Chain (each entry contains the SHA-256 of the previous entry and

includes a time-stamped hash); Add-Only Log Storage (write-once, no delete); Forensic Access (admin view only - complete attribution). All events in the system are captured, including user authentications, granted and denied access, file operations, the submission of AI queries and the AI-generated results, the generation of rationale reports, and actions taken by administrators. Hash chaining serves to prove the integrity of the log mathematically. Therefore, Layer 3 mitigates STRIDE threats R1 and R2 and also closes the NIST CSF PR.PT and DE.CM and ISO/IEC 27001 Control 8.15 gaps.

D. Layer 4 Ethical Governance and Compliance

Creating governance for all Automated Decision Making (ADM) operations begins with Layer 4 (Normative and Policy Envelope Governing Automated Decision-Making [ADM] Operationally). Layer 4 consists of five separate components: (1) An AI Security policy which benchmarks vendors' security before the AI platform's adoption (ISO 5.1, 5.19); (2) Implementing a Data Minimization protocol restricting the use of client data for AI processing to the minimum required as per GDPR Art. 5(1)(c); (3) Using Adversarial Input Validation to sanitize the input and verify the output (ISO 8.7); (4) Providing Cybersecurity Awareness Training (annual security training for all designers) to comply with role-specific security training for all design staff (ISO 6.3); and (5) Implementing Bias Auditing and Compliance Review using objective ethical review of AI systems based on AI4People and the EU AI Act. The completion of Layer 4 will also serve to mitigate the STRIDE threats T2, S3, and I3 and close the gaps related to NIST Cybersecurity Framework identification and risk assessment (ID.RA) and Role-appropriate security training (PR.AT), as well as ISO/IEC 27001 Controls 5.1, 5.7, 5.19, 6.3, and 8.7.

8. FRAMEWORK EVALUATION AGAINST STANDARDS

Table IX presents the post-ADI framework control coverage assessment, demonstrating measurable improvement across all assessed NIST CSF and ISO/IEC 27001 control domains.

TABLE IX: Pre- vs. Post-ADI Framework Control Coverage Assessment

Control Domain	**Standard Reference**	**Pre-ADI**	**Post-ADI**	**ADI Layer**	**STRIDE Threats Mitigated**
Identity & Access Mgmt	NIST PR.AC / ISO 8.2	**PM**	**FM**	L1: RBAC	S1, S2, S3, E1, E2, E3
Data Encryption	NIST PR.DS / ISO 8.24	**PM**	**FM**	L1: RBAC	I4, D2
AI Explainability	NIST PR.PT / EU AI Act 13	**NM**	**FM**	L2: XAI Engine	T1, R2, I2, I3
Audit Logging	NIST PR.PT / ISO 8.15	**NM**	**FM**	L3: Audit Infra.	R1, R2
Threat Assessment	NIST ID.RA / ISO 5.7	**NM**	**FM**	L4: Governance	All 18 threats
Security Awareness Training	NIST PR.AT / ISO 6.3	**NM**	**FM**	L4: Governance	S3

Adversarial Input Validation	NIST PR.PT / ISO 8.7	**NM**	**FM**	L2+L4	T1, T2
AI Vendor Security Policy	ISO 5.1, 5.19	**NM**	**FM**	L4: Governance	S2, T2, E3
Incident Response	NIST RS.RP	**NM**	**PM**	L4: Governance	I1, D1, D2
Anomaly / Behavioural Detection	NIST DE.AE	**NM**	**PM**	L3: Audit Infra.	I2, I3
COVERAGE: Pre-ADI: 0 FM, 4 PM, 6 NM (0% FM) → Post-ADI: 8 FM, 2 PM, 0 NM (80% FM) | Threat Mitigation: 13/13 Critical, 4/5 High					

9. DISCUSSION

According to the STRIDE and DREAD Threat Model, 72% of the identified threats fit into the Critical Risk classification, with the majority of High Severity Threats classified as Information Disclosure (I1: 8.2 and I4: 8.0) or Repudiation (R2: 8.4 and R1: 7.8) threats. The findings of the Gap Analysis explain the majority of these classifications; due to the absence of an Audit Infrastructure (NM, under both NIST PR.PT and ISO 8.15), and a lack of Explainability mechanisms in current Artificial Intelligence design platforms, conditions exist that allow for Repudiation attacks to occur with little to no possibility of prevention. The Gap Analysis identifies no NIST CSF Control Domain or ISO/IEC 27001 Control Domain as Fully Met under Baseline conditions, quantifies the extent of the governance deficit using empirical data, and provides a rigorous basis for creating an ADI framework. Methodologically, the empirical foundation of the threat model is a unique aspect to this research; the alignment of reported domains of high concern from survey data (data privacy [90%], auditability [58%], accountability [66%]) and the highest-severity STRIDE threats (I1: 8.2, R1/R2: 7.8/8.4) demonstrates that practitioner perceptions are appropriately calibrated to actual threat severity for this domain. This indicates that the ADI Framework adequately addresses real operational issues rather than merely offering a theoretical discussion of security. The fact that there's a significant relationship here is that there is statistically significant evidence of lower trust from senior professionals than more experienced professionals, and that there is lower trust in AI output when there is little explanation of how an AI generated that output ($\chi^2 = 7.42$, $p = 0.024$). This finding has implications for the design of the explanatory interface layer for XAI, in terms of providing detailed enough explanatory information (the explanation provided to senior professionals needs to give a very detailed level of technical detail through reports of the reasoning behind the AI's decision, input parameters, data source, and how confident the AI was in its result) rather than just a general explanation sufficient for a student user of the AI. Furthermore, this is consistent with extending the framework of trust in automated systems (Lee and See [15]) into a new field of creative professionals. Improvements to the ADI framework have resulted in a measurable, dramatic security enhancement, increasing Fully Met Control Coverage from 0% to 80%. There are only two remaining domains that are Partially Met: Incident Response and Anomaly Detection. These two domains are unavoidable due to the difficulties of implementing Behavioural Analytics and AI-based Incident Response Procedures in a resource-strained design firm environment. Therefore, they will be considered a top priority for the next iteration of the framework. The zero-trust architecture at the heart of the framework mandates explicit validation of every Access Request, regardless of the request's

source or Network. This is an intentional architectural decision that is appropriate for the multi-stakeholder, geographically dispersed nature of current design practice.

10. LIMITATIONS AND FUTURE WORK

A. Limitations

The statistical generalizability of Empirical Findings will be limited because of the survey Sample of (n=50) convenience sampling methods. The gap analysis relied on publicly available platform documentation rather than a direct Security Audit of the Platform infrastructure. Although the ADI Framework has been analytically compared to Security Standards, it has not yet been implemented in a live design environment, thus requiring further research to establish the use of the ADI Framework's empirical Measurement of Security Effectiveness and Workflow Friction, which would occur after its implementation. Triangulation of DREAD Scores should take place through future work by employing a Structured Expert Panel Review. Additionally, this study did not consider jurisdictional variations in Regulatory Requirements that lead to differences in governance responsibilities between countries.

B. Future Research Directions

The five main research priorities that can result from this project include: 1) an empirical evaluation of security effectiveness and adoption friction for the ADI Framework, specifically for the RBAC Engine and XAI Rationale Report Generator, through prototype implementation in one of the design firms participating in this study; 2) conducting a Delphi method expert panel review to provide and validate DREAD score triangulation; 3) performing formal penetration tests against the AI-based design platform under the governance of the ADI Framework to empirically test threat mitigation claims; 4) extending the threat model to adjacent creative professions also (e.g., architecture, fashion design and graphic design) to assess generalizability to AI-assisted creative workflows; and, 5) establishing longitudinal tracking of the evolution of the threat landscape as generative AI capabilities evolve as a means of preserving the contemporary relevance of the STRIDE/DREAD model.

CONCLUSION

This document is a cybersecurity threat model for AI-assisted interior design platforms, employing the STRIDE threat classification and DREAD risk quantification to identify, categorise, and rate specific threats associated with 18 well-defined cyberattacks. Results of the investigation demonstrated that 72% of the threats discovered were at a Critical risk severity. The greatest concentration of high-priority threats occurred within three categories: Information Disclosure, Repudiation, and Spoofing. A structured gap analysis benchmarking existing deployments of AI-based platforms against both NIST CSF v1.1 and ISO/IEC 27001:2023 determined that none of the control domains had been Fully Met, which results in an unqualified and quantifiable deficit in governance. As a result of this study, we propose the Augmented Design Intelligence (ADI) Framework: a four-layer cybersecurity governance framework through Adaptive RBAC Access Governance, XAI Transparency Engine, Immutable Hash-Chained Audit Infrastructure, Ethical AI Compliance Governance, and it is founded upon sound Security Engineering principles that researchers established during peer-reviewed studies on Cybersecurity. Our evaluation of the ADI Framework has demonstrated a significant increase in the percentage of Fully Met control coverage from 0% to 80%. It has mitigated all 13 Critical-rated STRIDE-assigned threats. The ADI framework provides a new, domain-specific way to govern AI platforms in interior design through a technically unique approach: it is the first standard-compliant cybersecurity governance architecture specifically created for AI-enabled interior design platforms. Regarding the specific design process of an interior designer’s project (i.e., via role-based access control, required XAI rationale reports,

and hash-chained audit trails for both humans and AI), the ADI framework is differentiated from both traditional IT security frameworks. It is a practical, deployable solution for designing firms of any size. Cybersecurity should be considered the necessary basis for using AI in creative design practises, not an impediment to it.

REFERENCES

[1] S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 4th ed. Hoboken, NJ, USA: Pearson, 2021.

[2] B. Shneiderman, "Human-centered artificial intelligence: Reliable, safe & trustworthy," Int. J. Human–Comput. Interact., vol. 36, no. 6, pp. 495–504, 2020.

[3] D. Gunning and D. Aha, "DARPA's explainable artificial intelligence (XAI) program," AI Mag., vol. 40, no. 2, pp. 44–58, 2019.

[4] F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," arXiv:1702.08608, 2017.

[5] I. Goodfellow, P. McDaniel, and N. Papernot, "Making machine learning robust against adversarial inputs," Commun. ACM, vol. 61, no. 7, pp. 56–66, 2018.

[6] A. Shostack, Threat Modelling: Designing for Security. Indianapolis, IN, USA: Wiley, 2014.

[7] International Organisation for Standardisation, ISO/IEC 27001:2023 Information Security Management Systems. Geneva, Switzerland: ISO, 2023.

[8] European Commission, "Regulation (EU) 2024/1689 Artificial Intelligence Act," Official J. EU, 2024.

[9] L. Floridi et al., "AI4People An ethical framework for a good AI society," Minds Mach., vol. 28, no. 4, pp. 689–707, 2018.

[10] H. Nissenbaum, Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford, CA, USA: Stanford Univ. Press, 2010.

[11] ENISA, Cloud Security for SMEs: Guidelines and Recommendations. Heraklion, Greece: ENISA, 2021.

[12] P. Mell and T. Grance, "The NIST definition of cloud computing," NIST Special Publication 800-145, 2011.

[13] National Institute of Standards and Technology, Framework for Improving Critical Infrastructure Cybersecurity, v1.1. Gaithersburg, MD, USA: NIST, 2018.

[14] B. Friedman and D. Hendry, Value Sensitive Design: Shaping Technology with Moral Imagination. Cambridge, MA, USA: MIT Press, 2019.

[15] J. Lee and K. A. See, "Trust in automation: Designing for appropriate reliance," Hum. Factors, vol. 46, no. 1, pp. 50–80, 2004.

[16] M. Boden, Creativity and Artificial Intelligence. Oxford, U.K.: Oxford Univ. Press, 2016.

[17] A. Bryman, Social Research Methods, 5th ed. Oxford, U.K.: Oxford Univ. Press, 2016.

[18] N. Papernot et al., "Technical challenges in machine learning security," IEEE Secur. Privacy, vol. 16, no. 3, pp. 55–62, 2018.

[19] R. H. Weber, "Internet of things n New security and privacy challenges," Comput. Law Secur. Rev., vol. 26, no. 1, pp. 23–30, 2010.

[20] World Economic Forum, Global Risks Report 2024. Geneva, Switzerland: WEF, 2024.

[21] L. Floridi et al., "An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations," Ethics Inf. Technol., vol. 22, pp. 689–707, 2018.

[22] S. Subashini and V. Kavitha, "A survey on security issues in service delivery models of cloud computing," J. Netw. Comput. Appl., vol. 34, no. 1, pp. 1–11, 2011.

[23] X. Yuan, P. He, Q. Zhu, and X. Li, "Adversarial examples: Attacks and defenses for deep learning," IEEE Trans. Neural Netw. Learn. Syst., vol. 30, no. 9, pp. 2805–2824, 2019.

[24] O. Sami Oubbati, A. Lakas, and F. Zhou, "A survey on position-based routing protocols for flying ad hoc networks (FANETs)," Veh Commun., vol. 10, pp. 29–56, 2017.

[25] M. Abomhara and G. M. Køien, "Cyber security and the internet of things: Vulnerabilities, threats, intruders and attacks," J. Cyber Secur. Mobility, vol. 4, no. 1, pp. 65–88, 2015.

[26] K. Stouffer, J. Falco, and K. Scarfone, "Guide to industrial control systems security," NIST Special Publication 800-82, Rev. 2, 2015.

[27] A. K. Sood and R. J. Enbody, "Targeted cyberattacks: A superset of advanced persistent threats," IEEE Secur. Privacy, vol. 11, no. 1, pp. 54–61, 2013.

[28] M. T. Khorshed, A. B. M. S. Ali, and S. A. Wasimi, "A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing," Future Gener. Comput. Syst., vol. 28, no. 6, pp. 833–851, 2012.

[29] A. Holzinger et al., "Explainable AI methods: A brief overview," in Machine Learning and Knowledge Extraction, Springer, 2022, pp. 13–38.

[30] Z. Obermeyer et al., "Dissecting racial bias in an algorithm used to manage the health of populations," Science, vol. 366, no. 6464, pp. 447–453, 2019.