ASTON UNIVERSITY

Aston Business School

MSc Cybersecurity Management

Defend as One: A Governance Framework for AI-Enhanced Automated Cyber Threat Intelligence in Public–Private Cyber Security Collaboration

PART 1: Research Context, General Literature Review and Conceptual Framework

Student name:

Student ID:

Supervisor:

Degree: MSc Cybersecurity Management

Submission: [

Table of Contents

List of Abbreviations

Abbreviation	Meaning
AI	Artificial intelligence
CAF	Cyber Assessment Framework
CiSP	Cyber Security Information Sharing Partnership
CTI	Cyber threat intelligence
DPA	Data Protection Act 2018
FIRST	Forum of Incident Response and Security Teams
GC3	Government Cyber Coordination Centre
IoC	Indicator of compromise
MISP	Malware Information Sharing Platform / MISP threat intelligence platform
MSSP	Managed security service provider
NCSC	National Cyber Security Centre
STIX	Structured Threat Information Expression
TAXII	Trusted Automated Exchange of Intelligence Information
TLP	Traffic Light Protocol
UK GDPR	United Kingdom General Data Protection Regulation

List of Tables

List of Figures

PART 1

RESEARCH CONTEXT AND LITERATURE REVIEW

Introduction • General Literature Review • Conceptual Framework and Research Gap

Chapter 1: Introduction

1.1 Background and Context

Cyber resilience in the UK nowadays depends on the organisational network in contrast to the technology estate in isolation. Several government agencies, local councils, critical national infrastructure owners, telecommunications companies, cloud and technology suppliers, managed security services providers (MSSPs), cybersecurity vendors, incident response suppliers and supply chain partners cooperate in multiple ways using services and depending on each other. In such a way, the threat to one environment can serve as an early warning to other environments if it is transformed into intelligence that can be understood by the receiving organisation. Unprocessed data or IoC is an observable artefact; tactical CTI gives insights about adversary's tactics; operational CTI helps in conducting campaigns and incident management; and strategic CTI is needed for risk and investment decision making. Hence, the actionable intelligence consists of context, confidence, and a decision-relevant recommendation (Tounsi and Rais, 2018; Ainslie et al., 2023). Thus, information exchange should be timely, structured, relevant, accurate, secure, and legal but not rapid and massive.

1.2 Cyber Threat Intelligence and Collaborative Cyber Defence

In collaborative cyber defense, an aim is to convert the detection of one party into prevention or reaction by others. Collaborative CTI helps enhance early warnings, correlation between organisations, prioritization of vulnerabilities, situational awareness and supply chain defense, along with avoiding redundant analysis (NIST, 2016; Wagner et al., 2019). Nevertheless, sharing does not always have positive impact. Redundant, outdated, irrelevant or low confidence indicators can lead to alert fatigue, misallocate limited resources of the analyst, and cause incorrect defensive actions to be taken. Therefore, the significance of collaboration lies in governance which will allow differentiating between urgent and validated intelligence and enable correction/retraction.

1.3 The UK 'Defend as One' Context

Cyber Resilience is defined in terms of being the responsibility of all the public sector through Government Cyber Security Strategy 2022-2030 and through the 'Defend as One' principle which ensures that there is coordination in detection, vulnerability management and incident response (Cabinet Office, 2022). The 2026 Government Cyber Action Plan maintains this course of action and names the Government Cyber Coordination Centre (GC3) as the operational mechanism which will be responsible for coordination of threats, vulnerabilities and incidents across the government (Cabinet Office, 2026; UK Government Security, 2026). It works alongside with other mechanisms which include the National

Cyber Security Centre (NCSC), Cyber Security Information Sharing Partnership (CiSP), Cyber Assessment Framework (CAF) and board level governance guidance (NCSC, 2025a; NCSC, 2025b; NCSC, n.d.). Although such mechanisms are vital, the difference among them needs to be understood as the strategy sets an ambition, GC3 deals with government operation coordination, CiSP offers a secure sharing community, while CAF and Board Toolkit deal with organisational assurance and oversight.

1.4 Organisational and Management Problem

The real challenge in this case is neither technical nor socio-technical but managerial. What is required is to make peace between speed and checking, between automation and control, between security and organisation interests, and between intelligence usefulness and secrecy. While public servants have their concerns regarding national resilience and statutory obligations, private companies also have concerns about protecting their customers, intellectual property, their reputation, and competitiveness. While technical standards may guarantee the machines' ability to read messages, they will not provide an answer to the question of whom these messages are going to, under which criterion, who is going to enrich these messages automatically, and how to address disputes regarding the indicators (Carr, 2016; Kotsias et al., 2023).

1.5 Research Problem and Rationale

Sharing of CTI by public and private organizations could be hindered on account of risks arising from reputation damage, legal ambiguity, risk of exposure of personal data, sensitivity of information, liability issues, lack of reciprocity, poor data quality and insufficient capacity for analysis. Research demonstrates that cooperation in CTI is as much determined by trusted relations and organizational context as technical proficiency (Zrahia, 2018; Piazza et al., 2023). Artificial intelligence-driven automation makes the issue more difficult because of factors like false positives, non-transparent decision-making, leakage of sensitive data and automation bias, all operating at machine speed.

1.6 Research Gap

The existing body of work has contributed useful insights that are nevertheless dispersed. Literature on technology has investigated STIX, TAXII, MISP, and interoperability; literature on privacy has proposed minimization and confidentiality measures; literature on organizations has addressed issues of trust, incentives, and public-private partnerships; literature on responsible AI has analyzed assurance and accountability. Yet the intersection of these lines of inquiry has hardly been investigated with regard to the issue of governing AI-enabled CTI automation in the UK (Preuveneers and Joosen, 2022; Alaeifar et al., 2024; Abraham et al., 2025). What needs to be clarified is how to have a meaningful framework

that covers trust/reciprocity, sensitive data, legal aspects, human control, accountability, data quality, interoperability, and auditability.

1.7 Research Question

How can public and private sector organisations improve trust, governance, and legal compliance in AI-enhanced automated cyber threat intelligence sharing?

1.8 Research Aim

The purpose is to explore the organisational, legal and trust-related challenges involved in AI-enhanced automated CTI sharing and propose a governance framework for secure, trusted, accountable and legal collaboration.

1.9 Research Objectives

1. Explore the contribution of AI-enhanced automated CTI sharing to cyber threat inference, detection, prioritization and incident response.
2. Explore the main barriers to public-private CTI sharing, including lack of trust, sensitivity of shared data, privacy issues, legal uncertainty, poor data quality and delayed sharing.
3. Explore the governance models, theories, standards and frameworks for secure and responsible cyber threat intelligence sharing.
4. Design a cyber security governance framework for AI-enhanced automated CTI sharing in UK public-private organisations.
5. Evaluate the framework from the perspectives of trust, privacy, accountability, interoperability and legal compliance.

1.10 Scope and Boundaries of the Study

The scope of the research includes UK public-private CTI sharing, organisational governance, AI-enhanced automation, trust, privacy, legal compliance, accountability, interoperability and human supervision. The research will not develop a new CTI platform, carry out penetration testing, gather live indicators, deconstruct malware, assess classified systems or acquire any confidential intelligence and vulnerabilities. The later qualitative phase will not aim at statistical generalisation to all UK organisations, but at understanding of the phenomenon.

1.11 Academic and Practical Significance

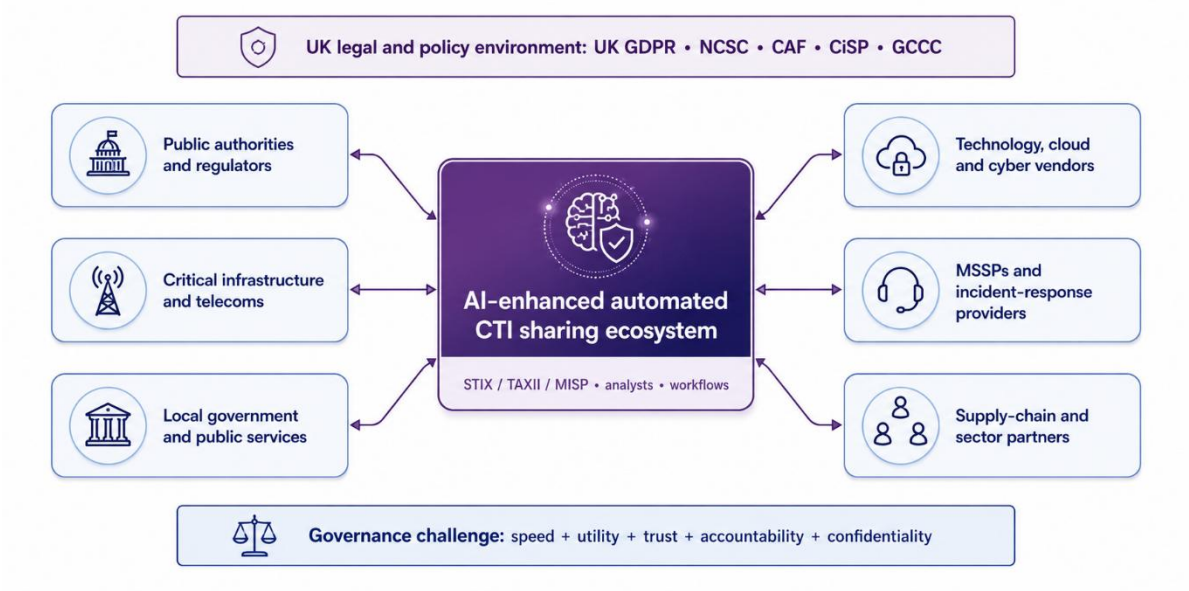
Academically, the research combines CTI with socio-technical systems, trust, information exchange, risk-governance and public-private partnerships perspectives. It also situates the responsible use of AI within the context of intelligence exchange instead of considering it an independent issue of compliance. Practically, the developed governance framework

will help with developing sharing agreements, classifying the sensitive data, defining access permissions, validating the shared information, maintaining audit trails, setting up human-review thresholds, reporting escalations and corrections. Such control mechanisms will allow to justify the trust rather than assume it and provide openness.

1.12 Structure of the Remaining Research Project

In Chapter 2, the literature will be critically reviewed and the literature-derived conceptual framework will be developed. In the future stages of research, methodology and ethical issues, data collection and thematic analysis, results, framework development and evaluation, discussion, conclusions and practical implications will be covered. No primary findings will be presented in Part 1.

Figure 1. Public–Private AI-Enhanced CTI Sharing Context



Source: Author's synthesis based on Cabinet Office (2022; 2026), NIST (2016), NCSC (2025a; 2025b) and Wagner et al. (2019).

Chapter 2: General Literature Review

2.1 Introduction to the Literature Review

This chapter brings together the technical, organisational, legal, and governance literature pertaining to CTI sharing via AI-powered automation. The discussion begins with CTI fundamentals and machine-readable exchange, moving on through public/private relations, privacy considerations, human control, and finally policy in the UK. This chapter does not attempt to document technology, but rather to understand the governance context within which CTI remains relevant and credible.

2.2 Cyber Threat Intelligence: Definitions, Forms and Lifecycle

CTI is not limited to threat information. Information is organization of observations like hashes, domains, logs or behavior, whereas data is the raw observation itself. Intelligence is the interpretation of observations in context of actor, capability, intention, target or decision process. NIST (2016) mentions IoCs, tactics, techniques and procedures, defensive measures and incident analysis as elements of cyber threat information. Academic research also indicates that CTI can only become actionable when it is relevant for certain decision (Shin and Lowry, 2020; Ainslie et al., 2023). Strategic CTI informs executive decision and policy making; operational CTI involves campaigns and incidents; tactical CTI provides rationale behind adversary's actions; and technical CTI is the quickly consumed observables. Lifecycle is typically connected with directions, collection, processing, analysis, dissemination and feedback (Tounsi and Rais, 2018). Lifecycle unveils the governance issues at each step: collection should be legal, processing should preserve provenance, analysis should convey confidence level, dissemination should comply with handling procedures, and feedback should fix or delete intelligence. Failure to implement these controls could make speed an authoritative-looking machine readable uncertainty.

Table 1. Key Definitions and Levels of CTI

Concept / level	Purpose	Typical content	Primary users	Governance concern
Raw threat data	Records observations	Logs, hashes, domains, alerts	Analysts, tools	Provenance, accuracy, retention
Technical CTI	Supports rapid detection	IoCs, signatures, infrastructure	SOC and incident-response teams	Expiry, false positives, machine ingestion
Tactical CTI	Explains adversary behaviour	Tactics, techniques and procedures	Defenders and threat hunters	Context, confidence, interpretation
Operational CTI	Supports campaign and incident decisions	Actor activity, targeting, timelines	Incident leaders and sector partners	Sensitivity, attribution, escalation

Concept / level	Purpose	Typical content	Primary users	Governance concern
Strategic CTI	Supports risk and policy	Trends, capability, intent, business impact	Boards, executives and policymakers	Uncertainty, proportionality, accountability
Actionable intelligence	Connects evidence to a decision	Context, confidence, recommended response	Decision owner	Explainability and ownership

Source: Author's synthesis based on NIST (2016), Tounsi and Rais (2018), Shin and Lowry (2020) and Ainslie et al. (2023).

2.3 Automated and Machine-Readable CTI Sharing: STIX, TAXII and MISP

The STIX provides standard representation for encoding threat actors, campaigns, indicators and other CTI objects, whereas TAXII protocol allows for the automated exchange of CTI objects between platforms (OASIS, 2021a; OASIS, 2021b). MISP can provide for collaboration, data modeling, distribution settings, and sharing groups (Wagner et al., 2016; MISP Project, n.d.). All these tools together can reduce the need for manual transcriptions, enable the automation of ingesting processes, and allow for correlations between different sources of CTI. Nevertheless, interoperability does not equate semantic interoperability. Variability of vocabularies, confidence scales and levels of granularity of objects, existence of duplicates and lack of context in indicators feeds, as well as access control configuration require certain competences (Dandurand and Serrano, 2013; Sauerwein et al., 2017). Traffic Light Protocol (TLP) may be used to indicate conditions under which the information may be redistributed; however, TLP does not replace legal assessments, agreements and verification of identity of the recipient (FIRST, 2022). Thus, there should be common data quality rules, metadata, handling guidelines and proper management of the platforms.

One more limitation is that the machine-readable formats do not imply absence of interpretative differences. The source of a particular STIX object, its analytical background, confidence scale, and domain of application may remain unknown although the object is technically valid. In such case, users will automatically ingest information without understanding of the threat level. There should be assessment of interoperability at syntactic, semantic, organizational, procedural and protocol levels in order to avoid a situation when technical interconnectivity becomes a fast channel for noise.

2.4 Artificial Intelligence in CTI Enrichment, Correlation and Prioritisation

AI could help with entity extraction from threat reports, indicator enrichment, clustering, cross-source correlation, classification, anomaly detection, confidence scoring, summarisation and recommendation support. The natural language processing would be able to detect IoCs and relationships between them automatically, while machine learning could detect IoCs that cannot be detected through manual analysis (Liao et al., 2016;

Alaeifar et al., 2024). Such capabilities differ from the deterministic rule-based automation because of probabilistic nature of the machine learning outputs and fluent and unsupported enrichments of the generative AI. Among benefits of the AI there are the speed and consistency of its operations as well as reduced analyst workload. At the same time, those techniques produce false positives and false negatives, suffer from model drift, adversarial manipulation, biases, contextual hallucination and overreliance on confidence scoring. Therefore, the UK AI Cyber Security Code of Practice requires the secure design and life-cycle risk management of the AI system, its monitoring and accountability (DSIT, 2025). In the case of CTI sharing, assurance should address the link between the performance of the model and its risks of dissemination. The low-confidence attribution, disclosure of the personal data or intelligence that may cause disruptions should be verified by humans, while routine high-confidence enrichments could be monitored in a human-on-the-loop mode. The audit logs should store the information about the source data, model version, applied transformations, confidence level, review decision and subsequent correction.

The quality of AI support depends on the quality of the evidence that has been used for the training, prompting and testing of the algorithm. The threat reporting is distributed unevenly across the industries and languages, while the historical labels could reproduce the biases and blind spots of the analysts. The model, which performs well for popular malware families, may prove to be unreliable in the cases of the novel campaigns or the different organisational environment. For that reason, the governance should require task-specific validation, uncertainty communication, drift monitoring and fall-back procedures. Explainability of the AI should depend on the use case: while analysts require evidence and provenance, executives require explanation of the automated recommendation.

Table 2. Benefits and Risks of AI-Enhanced Automated CTI Sharing

AI-enabled function	Potential benefit	Material risk	Governance response
Entity and IoC extraction	Processes large unstructured reports	Incorrect or decontextualised extraction	Validation rules and provenance
Enrichment and correlation	Connects dispersed evidence	False associations and sensitive-data exposure	Confidence thresholds and minimisation
Prioritisation and classification	Focuses analyst attention	Bias, model drift and automation bias	Performance monitoring and human challenge
Summarisation / generative AI	Reduces reading burden	Hallucinated facts or omitted caveats	Source-linked output and mandatory review
Recommendation support	Accelerates response planning	Disproportionate or disruptive action	Risk-tiered approval and escalation
Automated dissemination	Improves speed and scale	Inappropriate recipients or uncontrolled propagation	Access control, TLP/handling rules and audit logs

Source: Author's synthesis based on Liao et al. (2016), DSIT (2025), Alaeifar et al. (2024) and Abraham et al. (2025).

2.5 Public–Private Cyber Security Collaboration, Trust and Reciprocity

Government and industry are inter-dependent since important services, cloud computing, telecoms, managed security and significant incident telemetry all rest with private companies, whereas government agencies have at their disposal national-level intelligence and coordination mechanisms. The literature on public–private partnership cautions, however, that 'partnership' can mask important differences in authority, incentive structure, accountability and risk tolerance (Carr, 2016; Hodge and Greve, 2007). Major technology suppliers, minor suppliers, CI owners, regulators and law-enforcement agencies are quite different types and cannot be lumped together. Their capacity for intelligence gathering, analysis and sharing as well as their susceptibility to any negative effects of doing so varies considerably.

As a consequence, trust is multi-faceted. Competence-based trust concerns whether the partner in question is able to produce reliable CTI, integrity-based trust concerns its proper use, institutional trust comes from rules and oversight, technical trust relates to identity, access and system assurance, while procedural trust concerns consistent handling and correction. Trust is conceptualised by Mayer, Davis and Schoorman (1995) as willingness to rely on another party given his/her ability, benevolence and integrity. In turn, CTI studies show the relation between willingness to share and source reliability, reciprocity, reputation and controls against CTI misuse (Wagner et al., 2018). As shown by Zrahia (2018), there is the issue of coopetition since vendors compete among themselves but cooperate against common threat. Governance can reduce vulnerability but not eliminate it altogether.

Reciprocity is not necessarily about the equality in volume of contributions. Some participants can contribute unique context, some a large volume of telemetry data, and others coordination capacity. Thus, a healthy system requires clarity regarding what contributions can be expected from sources and information on how the intelligence contributed has been utilized, fixed or protected by the community. This will improve the reputation of the source and increase the learning process, but will also help with applying sanctions and revoking access rights to minimize abuse risk.

2.6 Organisational Incentives and Barriers to CTI Sharing

Technical accessibility does not necessarily translate into participation. There are fears among organizations about damaging their reputation, losing customers, facing investigations, exposing weaknesses, free riding, or falling behind in the competition. Resource-poor organizations might end up utilizing intelligence but not producing any since they do not have personnel to validate and sanitize it. On the other hand, mandatory reporting might increase the flow but not make the information more useful. There is evidence that factors such as support from managers, trusted dissemination mechanisms,

capabilities of practitioners, and organizational routines determine whether sharing becomes embedded in practice in the case of CTI uptake and cooperation between UK universities (Kotsias et al., 2023; Piazza et al., 2023).

2.7 Data Quality, Timeliness and Intelligence Utility

Intelligence that comes too late is not useful, but if intelligence is shared too early, then there would be an erroneous propagation of intelligence. The intelligence becomes good if it is relevant, accurate, complete, timely, contextualized, confident, trustworthy, from a credible source, provenance, and actionable (Wagner et al., 2019; Ainslie et al., 2023). Automation in intelligence sharing causes more of both the positive and negative impact because once the intelligence is incorrect, then it can be used by other users before it is validated. There must be a process put in place to have confidence scoring, source validation, de-duplication, versioning, expiration, revocation, and feedback.

2.8 Privacy, Confidentiality and UK Legal Compliance

The CTI can comprise IP addresses, e-mail addresses, usernames, identifying information about the device, records of employees or customers, descriptions of incidents, vulnerabilities information, and commercial information. However, not all indicators automatically fall into the personal data category depending on the possibility of individual identification in particular circumstances. If personal data processing occurs, the UK General Data Protection Regulation (UK GDPR) and the Data Protection Act 2018 require processing to be lawful, fair, and transparent, purpose limitation, data minimisation, accuracy, storage limitation, integrity, and confidentiality, accountability (UK Parliament, 2018; ICO, 2026). Artificial Intelligence-assisted processing may increase the risks of inferential, fair and opaque outcomes as well as inaccurate results (ICO, 2023).

Studies in privacy-preserving processing show that suppression, hashing, encryption, and controlled correlation can reduce the risks of disclosure, however, purely technical methods do not resolve the problems of purposes, access, and retention (Preuveneers and Joosen, 2022). Furthermore, governance framework for privacy-preserving processing should incorporate data classification, data minimization, data redaction, where appropriate, pseudonymization, encryption, differential access, retention rules, data-sharing agreements, logging, and a data protection impact assessment where processing involves a high risk. The legal interpretation is always context-dependent, and a framework should involve a requirement of assessment instead of the assumption that security purposes are more important than any other purposes.

However, the indicator can become more sensitive from the legal and operational perspective when changing contexts. An IP address of a malicious infrastructure can be a low-sensitivity indicator in one feed and a high-sensitivity indicator associated with an

employee, a customer, or an incident investigation when it will be used together with other information. Thus, minimization should be performed taking into account the potential enrichment rather than the indicator itself. Purpose and access should be re-evaluated when using CTI for attribution, employee decisions, or referral to the law enforcement.

2.9 Accountability, Explainability and Human Oversight

There are certain issues in connection with AI assisted collaboration which cannot be solved through the use of standards – who owns the results of automation, who confirms the enrichment process, who approves the distribution of the message and what is done in case of misuses and investigation of errors? “Human in the loop” is an approach requiring authorization prior to executing the action, “human on the loop” includes automation but with continuous control over it, and “human in command” implies an opportunity to control limits, to stop the action and take responsibility for it. It is difficult to prove the necessity of automation in case of insufficient trust, escalation across organizations, disclosure of personal information, conflicting intelligence or information necessitating a defensive disruption. It is necessary to have the possibility of challenge, logging the evidence, control over performance, separation of responsibilities and definition of thresholds for escalation in case of responsible control (DSIT, 2025; Abraham et al., 2025).

2.10 Cyber Risk Governance and Theoretical Foundations

Socio-technical theory lies behind the process, because CTI outputs emerge as a result of interactions between technology, humans, organisations and policies. What is socio-technical theory? According to Bostrom and Heinen (1977), optimisation of one technical subsystem without taking into account social system cannot take place. Thus, we can look at CTI process comprehensively but need to convert this theory into control systems. Trust theory can help us understand the willingness of people to be vulnerable, even if the level of trust depends on type of partner, data and context (Mayer et al., 1995). The literature on information sharing points out the importance of exchange and reciprocity, while the risk governance literature highlights proportionality, transparency, accountability and management of uncertainty (Renn, 2008). Lastly, public-private partnership theory describes interdependencies and incentives of parties but sometimes fails to provide operational details of CTI process (Carr, 2016). In other words, those theories are behind the creation of the framework where standards facilitate the exchange, governance defines rights and responsibilities and trust emerges due to competence and integrity and procedures.

The last but not least theory is institutional pressure. For instance, institutions can encourage people to join CTI because of regulators', clients' and professionals' expectations about good practices despite any immediate benefits. Such theoretical

approach helps to explain the reasons for the convergence, but it over-emphasises compliance and ignores informal practices. That is why responsible-AI principles are needed.

Table 3. Comparison of Key Theoretical Perspectives

Perspective	Core contribution	Strength for this study	Limitation	Framework support
Socio-technical systems	Joint optimisation of people, technology, process and structure	Explains why platform-only solutions fail	Broad without operational controls	All dimensions and organisational readiness
Trust theory	Willingness to accept vulnerability based on perceived trustworthiness	Explains participation and reliance	Trust varies by partner and intelligence type	Trust, reciprocity and challenge mechanisms
Information-sharing theory	Value, incentives, reciprocity and exchange rules	Links contribution to perceived benefit	May simplify intelligence sensitivity	Participation, feedback and incentives
Risk governance	Proportionality, uncertainty, accountability and control	Supports tiered review and escalation	Rigid controls may slow urgent sharing	Assurance, auditability and human oversight
Public-private partnership	Interdependence with different authority and incentives	Explains cross-sector coordination	Can understate technical operations	Roles, responsibilities and coordination

Source: Author's synthesis based on Bostrom and Heinen (1977), Mayer et al. (1995), Hodge and Greve (2007), Carr (2016) and Renn (2008).

2.11 UK Policy and Collaborative Mechanisms

In UK, cyber security is becoming increasingly understood as an issue of enterprise and ecosystem governance. While Government Cyber Security Strategy identified 'Defend as One' vision, Government Cyber Action Plan and GC3 present a more explicit coordination mechanism for government (Cabinet Office, 2022; Cabinet Office, 2026; UK Government Security, 2026). CAF 4.0 allows conducting cyber risk and resilience assessments of critical functions, NCSC Board Toolkit places responsibility on organizational leadership (NCSC, 2025a; NCSC, 2025b). CiSP is a platform where organizations from the UK can share threat intelligence (NCSC, n.d.). All these collaborative mechanisms allow articulating key responsibilities but, taken separately, do not provide an integrated approach to AI-enabled CTI enhancement and assessment, decision thresholds of data protection, correction and accountability. Moreover, some of the guidance provided is either voluntary or specific for certain sectors. The remaining question in terms of governance is how to integrate these mechanisms with the proper legal, organizational and AI assurance controls.

2.12 Literature Synthesis

From the literature analysis, it is clear that CTI is valuable for collective defense, machine-readability is a necessity, and trust, privacy, quality and accountability affect the willingness

to participate. However, the existing tension between the speed and veracity of information; openness and confidentiality; automation and oversight; reciprocity and difference in capacity; standardization and contextual approach; national resilience and commercial interests becomes obvious. Technical literature explains the process of intelligence transfer, organizational literature proves that this process might not happen, and legal and responsible-AI literature shows that unlimited transfer is unacceptable. Unexplored question is how to integrate all these aspects into a functional governance framework.

2.13 Research Gap and Conceptual Framework

There is a considerable amount of mature literature on CTI standards and platforms, empirical evidence of organizational usage and trust, as well as advanced privacy and AI assurance controls. At the same time, the aforementioned literature strands are not yet sufficiently integrated in the UK public-private scenario. Lack of evidence exists on how the participants can collaboratively govern AI-enabled automated CTI while ensuring intelligence utility, confidentiality, adherence to UK data protection rules, human accountability, machine-readability and fast correction. Hence, proposed literature-based conceptual framework includes seven interrelated dimensions: trust and reciprocity; privacy and sensitivity of data; compliance with law; accountability and oversight; quality and interoperability of data; automation assurance; public-private collaboration. Cross-cutting mechanisms of provenance, access control, auditability, human review, escalation, feedback and organizational readiness should be considered. Further empirical research will explore the perception of the above-mentioned dimensions by practitioners.

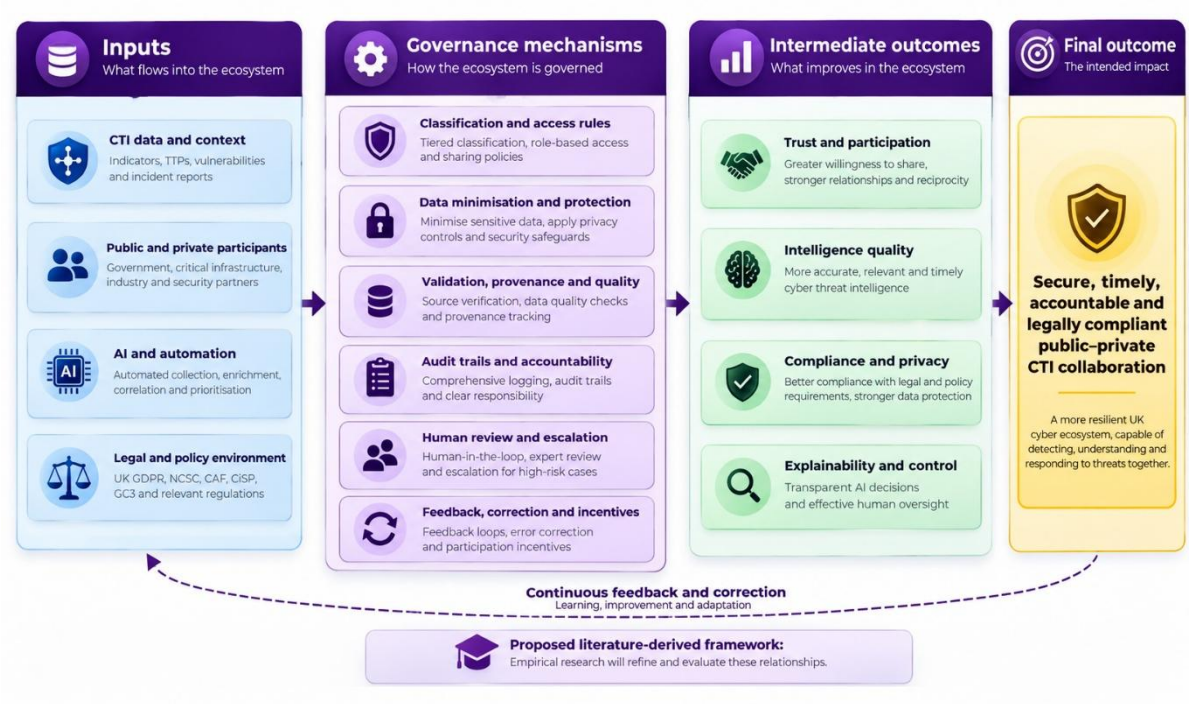
Table 4. Literature-Derived Governance Dimensions

Governance dimension	Core problem	Theoretical foundation	Proposed governance requirement	Expected outcome
Trust and reciprocity	Fear of misuse, free-riding and unreliable partners	Trust and information-sharing theory	Membership criteria, reciprocal feedback, sanctions and transparent handling	Justified confidence and sustained participation
Privacy and data sensitivity	Personal, confidential or commercially sensitive content	Socio-technical systems and privacy governance	Classification, minimisation, redaction, encryption and differential access	Proportionate disclosure and reduced harm
Legal compliance	Unclear basis, purpose, retention and accountability	Risk governance and institutional rules	Documented legal-basis assessment, sharing agreements, retention and DPIA triggers	Demonstrable compliance
Accountability and oversight	Unclear ownership of automated outputs and errors	Responsible AI and risk governance	Named decision owners, challenge routes, review thresholds and escalation	Traceable and contestable decisions

Governance dimension	Core problem	Theoretical foundation	Proposed governance requirement	Expected outcome
Data quality and interoperability	Duplicate, stale, inconsistent or context-poor intelligence	Information-quality and socio-technical perspectives	Provenance, confidence, deduplication, versioning, expiry and common schemas	Usable and reliable intelligence
Automation assurance	Model error, drift, bias and opaque transformation	Responsible AI and risk governance	Model/version logging, monitoring, validation and human-in-command controls	Safe, explainable automation
Public-private coordination	Different incentives, maturity, authority and resources	Public-private partnership and institutional theory	Clear roles, proportional contribution expectations and cross-sector escalation	Coordinated collective defence

Source: Author's synthesis based on the literature reviewed in Chapter 2.

Figure 2. Proposed Literature-Derived Conceptual Governance Framework



Source: Author's synthesis. The framework is literature-derived and has not been empirically validated.

2.14 Chapter Summary

From Chapter 2, it has been established that CTI sharing is a socio-technical exercise that requires governance. The theory explains the tensions in terms of the governance dimensions and forms the literature for the subsequent qualitative study.

References

Abraham, C., Bélanger, F. and Daultrey, S. (2025) 'Promoting research on cyber threat intelligence sharing in ecosystems', Journal of Cybersecurity, 11(1), tyaf016. doi: 10.1093/cybsec/tyaf016.

Ainslie, S., Thompson, D., Maynard, S. and Ahmad, A. (2023) 'Cyber-threat intelligence for security decision-making: A review and research agenda for practice', Computers & Security, 132, 103352. doi: 10.1016/j.cose.2023.103352.

Alaeifar, P., Pal, S., Jadidi, Z., Hussain, M. and Foo, E. (2024) 'Current approaches and future directions for cyber threat intelligence sharing: A survey', Journal of Information Security and Applications, 83, 103786. doi: 10.1016/j.jisa.2024.103786.

Bostrom, R.P. and Heinen, J.S. (1977) 'MIS problems and failures: A socio-technical perspective. Part I: The causes', MIS Quarterly, 1(3), pp. 17-32. doi: 10.2307/248710.

Cabinet Office (2022) Government Cyber Security Strategy: 2022 to 2030. London: HM Government. Available at: https://www.gov.uk/government/publications/government-cyber-security-strategy-2022-to-2030 (Accessed: 3 August 2026).

Cabinet Office (2026) Government Cyber Action Plan. London: HM Government. Available at: https://www.gov.uk/government/publications/government-cyber-action-plan (Accessed: 3 August 2026).

Carr, M. (2016) 'Public-private partnerships in national cyber-security strategies', International Affairs, 92(1), pp. 43-62. doi: 10.1111/1468-2346.12504.

Dandurand, L. and Serrano, O.S. (2013) 'Towards improved cyber security information sharing', in 5th International Conference on Cyber Conflict (CyCon 2013). Tallinn: NATO CCD COE Publications, pp. 1-16.

Department for Science, Innovation and Technology (DSIT) (2025) Code of Practice for the Cyber Security of AI. London: HM Government. Available at: https://www.gov.uk/government/publications/ai-cyber-security-code-of-practice (Accessed: 3 August 2026).

Forum of Incident Response and Security Teams (FIRST) (2022) Traffic Light Protocol (TLP): Version 2.0. Available at: https://www.first.org/tlp/ (Accessed: 3 August 2026).

Hodge, G.A. and Greve, C. (2007) 'Public-private partnerships: An international performance review', Public Administration Review, 67(3), pp. 545-558. doi: 10.1111/j.1540-6210.2007.00736.x.

Information Commissioner's Office (ICO) (2023) Guidance on AI and Data Protection. Available at: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/ (Accessed: 3 August 2026).

Information Commissioner's Office (ICO) (2026) A Guide to the Data Protection Principles. Available at: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-protection-principles/a-guide-to-the-data-protection-principles/ (Accessed: 3 August 2026).

Kotsias, J., Ahmad, A. and Scheepers, R. (2023) 'Adopting and integrating cyber-threat intelligence in a commercial organisation', European Journal of Information Systems, 32(1), pp. 35-51. doi: 10.1080/0960085X.2022.2088414.

Liao, X., Yuan, K., Wang, X., Li, Z., Xing, L. and Beyah, R. (2016) 'Acing the IOC game: Toward automatic discovery and analysis of open-source cyber threat intelligence', in Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, pp. 755-766. doi: 10.1145/2976749.2978310.

Mayer, R.C., Davis, J.H. and Schoorman, F.D. (1995) 'An integrative model of organizational trust', Academy of Management Review, 20(3), pp. 709-734. doi: 10.5465/amr.1995.9508080335.

MISP Project (n.d.) MISP Open Source Threat Intelligence Platform and Open Standards for Threat Intelligence Sharing. Available at: https://www.misp-project.org/ (Accessed: 3 August 2026).

National Cyber Security Centre (NCSC) (2025a) Cyber Assessment Framework 4.0. London: NCSC. Available at: https://www.ncsc.gov.uk/collection/cyber-assessment-framework (Accessed: 3 August 2026).

National Cyber Security Centre (NCSC) (2025b) Cyber Security Toolkit for Boards. London: NCSC. Available at: https://www.ncsc.gov.uk/collection/board-toolkit (Accessed: 3 August 2026).

National Cyber Security Centre (NCSC) (n.d.) Cyber Security Information Sharing Partnership (CiSP). Available at: https://www.ncsc.gov.uk/cisp/home (Accessed: 3 August 2026).

National Institute of Standards and Technology (NIST) (2016) Guide to Cyber Threat Information Sharing. NIST Special Publication 800-150. Gaithersburg, MD: NIST. doi: 10.6028/NIST.SP.800-150.

OASIS (2021a) STIX Version 2.1. OASIS Standard, 10 June. Available at: https://docs.oasis-open.org/cti/stix/v2.1/os/stix-v2.1-os.html (Accessed: 3 August 2026).

OASIS (2021b) TAXII Version 2.1. OASIS Standard, 10 June. Available at: https://docs.oasis-open.org/cti/taxii/v2.1/os/taxii-v2.1-os.html (Accessed: 3 August 2026).

Piazza, A., Vasudevan, S. and Carr, M. (2023) 'Cybersecurity in UK universities: Mapping (or managing) threat intelligence sharing within the higher education sector', Journal of Cybersecurity, 9(1), tyad019. doi: 10.1093/cybsec/tyad019.

Preuveneers, D. and Joosen, W. (2022) 'Privacy-preserving polyglot sharing and analysis of confidential cyber threat intelligence', in ARES '22: Proceedings of the 17th International Conference on Availability, Reliability and Security. New York: ACM, Article 2, pp. 1-11. doi: 10.1145/3538969.3538982.

Renn, O. (2008) Risk Governance: Coping with Uncertainty in a Complex World. London: Earthscan.

Sauerwein, C., Sillaber, C., Mussmann, A. and Breu, R. (2017) 'Threat intelligence sharing platforms: An exploratory study of software vendors and research perspectives', in Leimeister, J.M. and Brenner, W. (eds) Proceedings of the 13th International Conference on Wirtschaftsinformatik. St Gallen, pp. 837-851.

Shin, B. and Lowry, P.B. (2020) 'A review and theoretical explanation of the cyberthreat-intelligence (CTI) capability that needs to be fostered in information security practitioners and how this can be accomplished', Computers & Security, 92, 101761. doi: 10.1016/j.cose.2020.101761.

Tounsi, W. and Rais, H. (2018) 'A survey on technical threat intelligence in the age of sophisticated cyber attacks', Computers & Security, 72, pp. 212-233. doi: 10.1016/j.cose.2017.09.001.

UK Government Security (2026) Government Cyber Coordination Centre (GC3). Available at: https://www.security.gov.uk/services-resources/government-cyber-coordination-centre/ (Accessed: 3 August 2026).

UK Parliament (2018) Data Protection Act 2018, c. 12. London: The Stationery Office. Available at: https://www.legislation.gov.uk/ukpga/2018/12/contents (Accessed: 3 August 2026).

Wagner, C., Dulaunoy, A., Wagener, G. and Iklody, A. (2016) 'MISP: The design and implementation of a collaborative threat intelligence sharing platform', in Proceedings of the 2016 ACM Workshop on Information Sharing and Collaborative Security. New York: ACM, pp. 49-56. doi: 10.1145/2994539.2994542.

Wagner, T.D., Mahbub, K., Palomar, E. and Abdallah, A.E. (2019) 'Cyber threat intelligence sharing: Survey and research directions', Computers & Security, 87, 101589. doi: 10.1016/j.cose.2019.101589.

Wagner, T.D., Palomar, E., Mahbub, K. and Abdallah, A.E. (2018) 'A novel trust taxonomy for shared cyber threat intelligence', Security and Communication Networks, 2018, 9634507. doi: 10.1155/2018/9634507.

Zrahia, A. (2018) 'Threat intelligence sharing between cybersecurity vendors: Network, dyadic, and agent views', Journal of Cybersecurity, 4(1), tyy008. doi: 10.1093/cybsec/tyy008.