Article

The Technical–Regulatory Correspondence Matrix: A Practical Development Framework for Building GDPR- and AI Act-Compliant High-Risk AI Systems

Antonio Goncalves * and Anacleto Correia *

Posted Date: 8 December 2025

doi: 10.20944/preprints202512.0593.v1

Keywords: cybersecurity; privacy; EU AI Act; GDPR; high-risk AI systems; critical infrastructure protection; network anomaly detection; AI governance; regulatory compliance; MLOps; observability; evidence-by-design; auditability

 doi:10.20944/preprints202512.0593.v1

Article

The Technical–Regulatory Correspondence Matrix: A Practical Development Framework for Building GDPR- and AI Act-Compliant High-Risk AI Systems

Antonio Goncalves * and Anacleto Correia

CINAV; 2810-001 Almada;
* Correspondence: agoncalvesLX@gmail.com

Abstract

The European Union Artificial Intelligence Act (AI Act) and the General Data Protection Regulation (GDPR) impose stringent and partly overlapping obligations on high-risk AI systems deployed in cybersecurity and critical infrastructure contexts. Yet organisations still lack concrete mechanisms to translate these legal requirements into actionable engineering tasks and auditable evidence along MLOps lifecycles. This paper proposes the Technical–Regulatory Correspondence Matrix (TRCM) as a structured correspondence layer that explicitly links regulatory pillars (derived from the GDPR, the AI Act and emerging AI management system standards) to families of technical dimensions in AI-based security monitoring and incident detection. The TRCM captures the many-to-many relationships between legal obligations and technical activities and is designed to be instantiated for specific high-risk use-case families. We introduce the matrix, define its regulatory and technical dimensions, and apply it to a representative cybersecurity scenario: network anomaly detection operated by essential service operators to protect critical infrastructures. For this use case, we derive a regulatory profile, construct a filtered TRCM and show how obligations on risk management, data governance, robustness, transparency and human oversight can be mapped to concrete controls (for example, data inventories and lineage, stress-testing suites, monitoring and incident response procedures, explainability mechanisms and human–AI interaction patterns) and to associated evidence artefacts embedded as correspondence checkpoints in an MLOps pipeline. We then analyse the operational implications of adopting the TRCM for engineering, compliance, risk and audit functions, arguing that it supports an evidence-by-design posture and observability-driven AI governance in cybersecurity operations. Finally, we discuss the limitations of the current formulation and outline directions for future work on standardisation, automation, handling regulatory tensions between the AI Act and the GDPR, and multi-stakeholder deployments of the TRCM in network security and critical infrastructure ecosystems.

Keywords: cybersecurity; privacy; EU AI Act; GDPR; high-risk AI systems; critical infrastructure protection; network anomaly detection; AI governance; regulatory compliance; MLOps; observability; evidence-by-design; auditability

1. Introduction

1.1. Motivation and Problem Statement

Responsible artificial Intelligence (AI) development increasingly relies on frameworks that connect normative principles, legal requirements, and engineering practices across the entire system lifecycle [1, 2]. The rapid diffusion of AI in sensitive and high-stakes domains has intensified demands for transparency, accountability, and meaningful human oversight throughout the development and deployment of AI systems [1,3]. Conceptual frameworks for AI regulation have been proposed that span multiple stages of the policy process, but these contributions largely remain at a macro-governance

 doi:10.20944/preprints202512.0593.v1

level and offer limited guidance on how technical teams should translate regulatory expectations into concrete, verifiable controls and artefacts [1,2]. In parallel, analyses of AI politics in the European Union show that regulatory responses combine market-integration objectives with the reinforcement of fundamental-rights protection, thereby raising the bar for accountability and oversight, particularly for high-risk AI systems [4].

AI governance is increasingly characterised as a multidimensional and fragmented field, marked by coordination challenges, power asymmetries, institutional capacity gaps, and overlapping or competing regulatory initiatives [2,4]. Synthesis work highlights a proliferation of governance strategies, tools, and frameworks, while simultaneously stressing the lack of operational mechanisms that link these instruments to concrete development and deployment practices [2,5]. This disconnect creates a persistent gap between high-level narratives of "responsible" or "trustworthy" AI governance and the everyday work of engineers who must design metrics, pipelines, and technical documentation that satisfy specific obligations imposed by regimes such as the GDPR and the EU AI Act [2,3]. In practice, developers are left without practical guidance on how to systematically translate these regulatory requirements into verifiable technical artefacts that can be inspected and audited.

Recent research on responsible AI governance systematises the organisational structures, processes, and capabilities needed to operationalise values such as transparency, accountability, and human oversight [5]. This work distinguishes structural, relational, and process-oriented aspects of responsible AI governance, yet stops short of specifying how these dimensions should be instantiated as technical requirements, Machine Learning Operations (MLOps) controls, and auditable evidence in concrete AI projects [3,5]. Broader governance-of-AI analyses reinforce that the central challenge is no longer the absence of principles, but the difficulty of operationalising legal notions such as explainability, documentation, risk mitigation, and traceability into metrics, controls, and evidence integrated into development pipelines, in rapidly evolving socio-technical environments [2,4].

Within the trustworthy AI literature, integrative contributions connect legal, ethical, and technical robustness pillars, mapping requirements such as human oversight, transparency, privacy, and accountability and exploring how they relate to responsible AI systems and emerging regulation [3]. However, these efforts remain largely conceptual and do not deliver a systematic model that, for each legal obligation, specifies the corresponding metrics, controls, logs, reports, and development artefacts that should be produced and maintained throughout the lifecycle of a high-risk AI system [3]. From an MLOps perspective, there is still no unified framework that helps technical teams design and operate high-risk AI systems in a way that is natively compliant with emerging regulation. Taken together, these strands of work motivate the need for a technical–regulatory correspondence model and define the problem this article addresses: current AI regulation and governance frameworks do not yet provide an operational mapping that developers can embed into their MLOps processes to natively produce demonstrable compliance with instruments such as the GDPR and the EU AI Act [1–3,5].

1.2. Regulatory Landscape: GDPR and AI Act

The regulatory landscape for high-risk AI in the European Union is primarily shaped by the General Data Protection Regulation (GDPR) and the EU AI Act, which together establish complementary but distinct sets of obligations for organisations deploying AI systems that process personal data or affect individuals in significant ways [1,3,4]. While the GDPR focuses on data protection principles, individual rights, and accountability duties for controllers and processors, the AI Act introduces a risk-based framework that directly targets AI systems, imposing specific technical and organisational requirements on providers and deployers of high-risk AI [2,3]. Both instruments converge on core values such as transparency, fairness, human oversight, and robustness, but they operationalise these values through different legal mechanisms, documentation duties, and enforcement tools [1,4].

Under the GDPR, automated decision-making and profiling are subject to stringent transparency and accountability requirements, particularly where decisions produce legal or similarly significant effects on individuals [3]. Controllers must inform data subjects when automated decision-making is used, provide meaningful information about the logic involved, and ensure that individuals can

 doi:10.20944/preprints202512.0593.v1

contest decisions and obtain human intervention, while implementing appropriate safeguards to protect fairness, accuracy, and data protection rights [1,3]. These obligations translate into the need for clear documentation of processing activities, traceable decision paths, and technical and organisational measures that support effective oversight and redress.

The AI Act complements this framework by defining high-risk AI systems and subjecting them to a prescriptive set of requirements that span risk management, data governance, technical documentation, logging, transparency, human oversight, robustness, accuracy, and cybersecurity [2,3]. Providers of high-risk systems must implement a documented quality management system, conduct conformity assessments, and register systems in a dedicated EU database, while ensuring that event logging and record-keeping enable ex post analysis of system behaviour and incident reconstruction [1,4]. Deployers, in turn, are required to operate these systems in accordance with the provider's instructions, monitor performance, and maintain appropriate human oversight and documentation to manage risks and demonstrate ongoing compliance.

Taken together, the GDPR and the AI Act push organisations towards AI systems that are not only legally compliant but also explainable, traceable, and auditable by design [3]. However, neither instrument specifies in a systematic way how obligations such as transparency, human oversight, risk management, or logging should be instantiated as concrete metrics, controls, pipelines, and development artefacts in MLOps workflows for high-risk AI [1,2]. For development teams, this combined framework effectively requires the systematic production of verifiable, audit-ready technical artefacts across design, training, validation, deployment, and operation phases, while still leaving open the question of how to architect such artefacts in practice [3].

1.3. Objectives and Contributions

Building on prior work that maps AI principles and ethics to responsible AI systems and regulation [6,7] and on emerging frameworks and pattern catalogues for responsible AI governance and engineering [8,9], this article proposes a practical, development-oriented framework to support the engineering of high-risk AI systems in native alignment with the GDPR and the EU AI Act. The central artefact of this framework is the Technical–Regulatory Correspondence Matrix (TRCM), which organises selected legal obligations into a structured set of dimensions that can be directly mapped onto technical requirements, engineering controls, and MLOps artefacts across the AI lifecycle. Rather than treating governance and compliance as ex post verification activities, the TRCM is designed to guide design, implementation, and operation decisions so that regulatory expectations are embedded into the system from the outset.

The first objective of the article is to introduce and formalise the TRCM as a reusable model that relates concrete legal obligations to families of technical metrics, controls, and documentation patterns that can be instantiated in different high-risk AI contexts. The matrix captures how requirements such as transparency, human oversight, risk management, data protection, and logging can be expressed as verifiable properties of models, pipelines, and operational processes, enabling the systematic generation of compliance evidence throughout design, training, validation, deployment, and monitoring. This systematic correspondence aims to reduce the current gap between abstract regulatory texts and the everyday engineering work of development teams that must navigate complex governance demands across domains such as healthcare, finance, and public services [10–12].

The second objective is to specify a set of principles, technical dimensions, and mapping mechanisms that make the TRCM operational for practitioners. The framework delineates core dimensions such as data governance, model behaviour, explanation and communication, monitoring and incident response, and documentation and auditability, and shows how each dimension can be populated with concrete metrics, controls, and artefacts that correspond to individual legal obligations. By providing worked examples and implementation patterns, the article illustrates how regulatory requirements can be transformed into consistent, replicable engineering decisions, supporting both internal governance processes and external audits of high-risk AI systems [13,14].

 doi:10.20944/preprints202512.0593.v1

Finally, the article consolidates these elements into a coherent XAI-Compliance-by-Design approach that integrates legal, organisational, and technical perspectives on responsible AI [5,9]. The proposed framework is intended to be sufficiently general to apply across high-risk AI domains, while being concrete enough to support practical implementation in MLOps workflows. In doing so, it contributes a technical–regulatory correspondence model that helps organisations move from principle-level commitments to demonstrable, audit-ready compliance artefacts embedded in the lifecycle of AI systems.

1.4. Article Structure

The remainder of this article is organised as follows. Section 2 reviews the relevant literature on explainable AI (XAI), AI governance, and regulatory compliance, drawing on recent analyses that highlight the persistent gap between high-level normative principles and actionable technical guidance. This section also synthesises emerging work on socio-technical alignment, risk management, and observability-driven governance, emphasising the absence of operational mechanisms capable of translating legal obligations into verifiable engineering artefacts.

Section 3 introduces the Technical–Regulatory Correspondence Matrix (TRCM), describing its conceptual foundations, regulatory pillars, and technical dimensions. It further details the methodology used to construct the matrix and situates the proposal in relation to existing frameworks for transparency, accountability, documentation, and human oversight under both the GDPR and the EU AI Act.

Section 4 applies the TRCM to a representative high-risk AI scenario, demonstrating how legal requirements—particularly those related to data governance, risk management, robustness, transparency, and logging— can be mapped to concrete metrics, controls, and development artefacts. This section illustrates how the TRCM supports structured evidence generation and regulatory readiness throughout the AI lifecycle.

Section 5 discusses the broader implications of the TRCM for engineering practice, compliance processes, regulatory harmonisation, and multi-stakeholder coordination. It also examines limitations of the current formulation and identifies open research challenges.

Finally, Section 6 summarises the main contributions of the paper and outlines directions for future work, including the extension of the TRCM to additional AI system classes, sector-specific risk profiles, and complementary regulatory instruments.

2. Background and Related Work

2.1. Explainability, Governance and Compliance

Research on Explainable AI (XAI) consistently highlights that explainability, auditability, and meaningful human oversight are essential properties for high-risk AI systems deployed in sensitive domains such as healthcare, finance, and critical infrastructure [15–17]. Surveys and taxonomies show a rich variety of methods—ranging from feature-attribution techniques to counterfactual explanations and user-centred evaluation studies—yet these contributions typically frame explainability as a property of individual models or interfaces, rather than as a set of requirements that must be systematically integrated into end-to-end development and deployment processes [16,18,19]. As a result, many high-risk AI projects still treat XAI as an add-on layer applied late in the lifecycle, instead of as a first-class design constraint that shapes data, modelling, and operational choices from the outset.

Despite the progress in XAI methods, significant gaps remain regarding how to transform techniques such as SHAP, LIME, or counterfactual explanations into artefacts that can be used to demonstrate regulatory compliance in a consistent and repeatable way [15,17]. Existing studies rarely specify how explanation outputs should be logged, versioned, and linked to concrete decisions, or how they should be aligned with legal notions such as transparency, contestability, and meaningful information about the logic involved in automated processing. In practice, development teams are left without clear guidance on how to embed explanation mechanisms into MLOps pipelines so that they systematically

 doi:10.20944/preprints202512.0593.v1

generate evidence that is fit for purpose in audits, investigations, or supervisory reviews under regimes such as the GDPR and the EU AI Act [18,19].

Work on algorithmic governance and responsible AI further emphasises that technical explanation metrics alone are insufficient, and that governance structures must harmonise these metrics with legal and organisational requirements [13,14]. Reviews and frameworks in AI governance identify the need for integrated approaches that connect model performance, robustness, and explainability with risk management processes, documentation practices, and accountability mechanisms at organisational level [5,9]. However, these contributions typically stop short of providing operational instruments that specify how particular explanation methods, logs, and documentation artefacts should correspond to specific legal obligations, leaving a persistent gap between high-level governance principles and the concrete engineering work required to build audit-ready high-risk AI systems [10–12].

2.2. Existing Frameworks and Their Limitations

A variety of documentation-centric approaches have been proposed to improve transparency and accountability in machine learning, including documentation artefacts such as Model Cards and Datasheets for Datasets, which provide structured templates to describe the purpose, data provenance, evaluation metrics, limitations, and risks of models and datasets [16**?**]. These initiatives offer valuable structural guidance for documenting technical systems and are increasingly recognised as useful tools for internal governance and external communication in high-stakes AI applications [10,11]. However, they do not by themselves establish a direct correspondence between specific documentation fields or metrics and concrete legal obligations, nor do they prescribe how such artefacts should be integrated into development workflows to support continuous, audit-ready compliance with instruments such as the GDPR and the EU AI Act [6].

Beyond documentation, a growing body of work on algorithmic auditing, AI governance frameworks, and responsible AI pattern catalogues proposes checklists, assessment procedures, and best practices for evaluating AI systems after they have been designed or deployed [8,9]. These frameworks are predominantly evaluative and reactive, focusing on ex post reviews of models, identification of risks, and assessment of alignment with high-level principles or sectoral standards, and some offer domain-specific guidance for areas such as healthcare and medical imaging [10–12]. While useful for audits, procurement decisions, and oversight, they offer limited support as operational guides for engineering teams during the initial design and implementation of high-risk AI systems.

Taken together, existing documentation schemes, auditing frameworks, and governance models do not yet provide a unified methodology that links documentation artefacts, technical metrics, and MLOps controls to well-defined regulatory requirements in a systematic way [5,6]. Development teams still lack an instrument that tells them which logs, explanations, validation reports, and monitoring dashboards should be produced for each legal obligation, and how to embed these artefacts into pipelines so that compliance evidence is generated by design rather than assembled ad hoc. This gap motivates the need for a structured technical–regulatory correspondence model such as the TRCM, which explicitly connects regulatory obligations to families of technical artefacts across the AI lifecycle.

2.3. Need for a Unified Technical–Regulatory Mapping

Despite the abundance of technical methods, documentation schemes, and regulatory frameworks for AI, there is still no clear mechanism that systematically maps engineering artefacts to concrete requirements under the GDPR and the EU AI Act [5,6]. Existing work tends to treat legal and governance obligations as external constraints to be interpreted by lawyers or compliance teams, while leaving development teams to decide, largely ad hoc, which metrics, logs, reports, and documentation elements are sufficient to demonstrate compliance in high-risk AI contexts [9,11]. This disconnect makes it difficult to ensure that technical decisions taken during data preparation, model design, validation, deployment, and monitoring are traceably linked to specific legal rights, principles, and obligations.

 doi:10.20944/preprints202512.0593.v1

From an engineering perspective, the lack of a unified mapping is particularly problematic in MLOps settings, where pipelines, model registries, and monitoring infrastructures need to be configured in ways that continuously produce audit-ready evidence [8,10]. Without a structured reference model, teams must decide which events to log, which explanation artefacts to retain, how to structure documentation, and how to align these artefacts with regulatory expectations regarding transparency, human oversight, risk management, and robustness [1,2]. This situation increases the risk of both under-compliance (missing critical evidence) and over-compliance (producing redundant artefacts that are costly to maintain and difficult to interpret).

The Technical–Regulatory Correspondence Matrix (TRCM) introduced in this article is designed to address this gap by providing a formal instrument that aligns technical decisions with legal obligations from the design phase onwards [6,20]. The TRCM explicitly relates selected GDPR and EU AI Act requirements to families of metrics, controls, logs, reports, and documentation patterns that can be embedded into MLOps workflows for high-risk AI systems. By making these correspondences explicit, the TRCM enables compliance to be engineered by design rather than retrofitted ex post, supporting development teams in systematically producing verifiable, audit-ready artefacts across the AI lifecycle.

3. The Technical–Regulatory Correspondence Matrix

3.1. Conceptual Overview

The Technical–Regulatory Correspondence Matrix (TRCM) is proposed as a development-oriented framework that closes the technical–regulatory gap between the legal obligations defined in the GDPR and the EU AI Act and the concrete artefacts produced in the engineering of high-risk AI systems, as summarised in Figure 1.

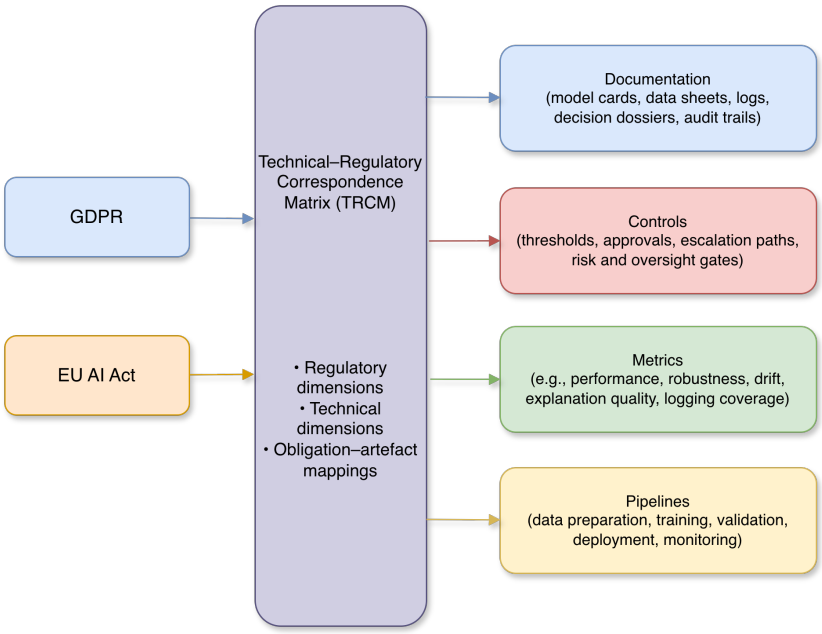

Figure 1. High-level conceptual structure of the TRCM.

Rather than treating regulation as an external checklist applied after models are built, the TRCM embeds regulatory requirements directly into design, implementation, and operation decisions, so that compliance, explainability, and auditability become structural properties of the system rather than late-stage add-ons.[5,6] In this sense, the TRCM responds to a central weakness identified in the AI governance and trustworthy AI literature: the absence of a systematic, operational bridge that links abstract legal duties to the metrics, pipelines, controls, and documentation that govern day-to-day engineering practice.[8,9]

Conceptually, the TRCM organises legal obligations into a set of regulatory and technical dimensions that capture the main areas in which compliance must be demonstrated in high-risk AI

 doi:10.20944/preprints202512.0593.v1

systems: data governance, model behaviour and performance, transparency and explainability, human oversight and contestability, risk management and robustness, logging and monitoring, and documentation and auditability.[6,11] Each dimension aggregates related GDPR and AI Act requirements and expresses them as verifiable properties that can be realised through specific technical artefacts and processes in MLOps workflows, such as dataset documentation, evaluation reports, explanation logs, and incident records.[1,20] By structuring obligations in this way, the TRCM provides a coherent scaffold that helps teams reason about how rights, principles, and duties should manifest in concrete artefacts spanning datasets, models, logs, reports, and decision dossiers.[5,10]

At the core of the TRCM lies a set of structuring principles that connect each mapped obligation to four families of technical elements: metrics, pipelines, controls, and documentation.[8,21] Metrics make compliance-relevant properties observable and quantifiable by, for example, capturing performance under distribution shift, robustness to perturbations, explanation stability, or logging coverage.[16,17] Pipelines embed these metrics into data preparation, training, validation, deployment, and monitoring workflows, so that evidence is generated continuously as part of normal operation rather than through ad hoc, episodic checks.[8,10] Controls implement thresholds, approvals, segregation of duties, and escalation mechanisms that operationalise obligations related to risk management, human oversight, and incident response.[9,13] Finally, documentation artefacts consolidate this evidence into structured, reusable forms—such as model and dataset cards, DPIA summaries, audit logs, and decision dossiers—that support both internal governance and external supervision under the GDPR and the AI Act.[6,22]

By articulating these dimensions and elements within a single correspondence matrix, the TRCM goes beyond prior governance and responsible AI frameworks that remain predominantly narrative, principle-based, or policy-oriented.[5,23] The matrix acts as a reusable blueprint that can be instantiated across domains and organisational contexts: for each selected obligation, it indicates which technical artefacts must exist, how they should be integrated into MLOps pipelines, and how they collectively form an evidence base traceable to specific GDPR and AI Act provisions.[20,24] This conceptual structure underpins the more detailed regulatory and technical dimensions developed in the following subsections, where the TRCM is refined into concrete mappings between obligations, metrics, controls, and documentation patterns that enable compliance to be engineered by design throughout the AI lifecycle.[6,8,9]

3.2. Matrix Construction Methodology

The construction of the TRCM follows a methodological process that starts from a systematic collection and structuring of legal requirements under the GDPR and the EU AI Act and ends with a reusable matrix that can be instantiated in concrete MLOps pipelines, as summarised in Figure 2.

 doi:10.20944/preprints202512.0593.v1

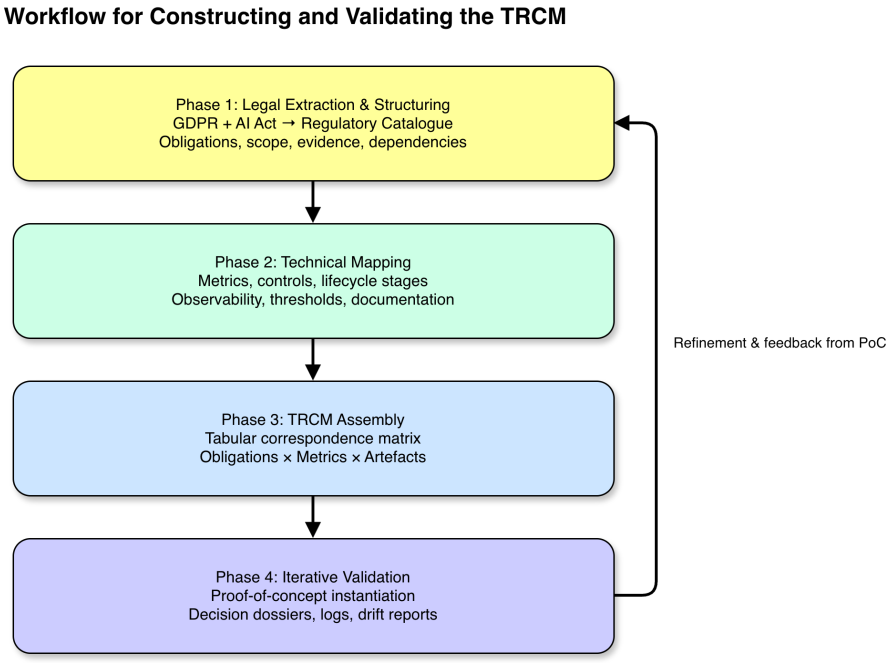

Figure 2. Overview of the TRCM construction methodology.

The aim is not only to catalogue obligations, but to express them as verifiable properties that can be operationalised through metrics, controls, and documentation patterns across the lifecycle of high-risk AI systems.[5,6] This process is designed so that developers and governance teams can use the TRCM as a practical design-time and run-time instrument, rather than as a purely conceptual compliance checklist.[8,9]

The first phase consists of extracting and organising relevant legal obligations from the GDPR and the AI Act, focusing on provisions that give rise to technical or documentation duties in high-risk AI contexts.[1,20] Obligations are grouped into the regulatory dimensions introduced in the previous subsection (data protection and data governance, risk management and robustness, transparency and human oversight, documentation and logging, post-market monitoring and incident management), taking into account how these duties are interpreted in the AI governance and trustworthy AI literature.[3,6] For each obligation, the methodology clarifies its scope (e.g., provider vs. deployer, design vs. operation), the type of evidence it requires, and its dependencies with other obligations, producing a structured regulatory catalogue that serves as the backbone of the matrix.[2,11]

In the second phase, this regulatory catalogue is systematically mapped to families of technical metrics, controls, and artefacts, drawing on existing work on responsible AI engineering, MLOps, and documentation practices.[8,16] For each obligation–dimension pair, the methodology identifies: (i) metrics that make the obligation observable (e.g., performance, robustness, drift, explanation quality, logging coverage); (ii) pipeline stages where these metrics must be computed (data preparation, training, validation, deployment, monitoring); (iii) controls and gates that enforce thresholds, approvals, or escalation paths; and (iv) documentation artefacts that capture the resulting evidence in a form suitable for audits and conformity assessments.[10,22] This step ensures that every selected obligation in the regulatory catalogue is associated with at least one concrete technical handle that can be embedded into development workflows.[6,8]

The third phase assembles these correspondences into the matrix structure itself, creating a tabular representation where rows represent individual or grouped legal obligations and columns represent technical elements such as metrics, logging and monitoring artefacts, documentation patterns, and governance processes.[6,9] Each cell in the matrix specifies one or more concrete artefacts (for example, a particular log, report, or explanation bundle) that should exist if the corresponding obligation is to be demonstrably satisfied, together with indications of responsible roles and lifecycle phases.[5,24] This representation makes it straightforward to derive implementation checklists, pipeline configurations,

 doi:10.20944/preprints202512.0593.v1

and evidence generation scripts for a given high-risk use case, while also supporting top-down views for auditors and regulators.[20,23]

Finally, the methodology incorporates an iterative validation loop in which the TRCM is instantiated in a reference high-risk AI scenario and refined based on implementation experience.[10,12] In this work, the anomaly detection proof of concept described in Section **??** is used to test whether the proposed correspondences can be realised in code, integrated into MLOps pipelines, and produce stable, audit-ready artefacts such as decision dossiers, explanation logs, drift reports, and documentation bundles.[8,13] Feedback from this instantiation is used to adjust the granularity of obligations, refine metric definitions, and simplify documentation patterns, with the goal of providing a matrix that is both faithful to the regulatory texts and practically usable by development teams from initial design through deployment and post-market monitoring.[5,6]

3.3. Regulatory Dimensions

The regulatory dimensions of the TRCM define the main legal pillars that must be reflected in the design, deployment and governance of high-risk AI systems. Instead of working with a long list of dispersed clauses in the GDPR and the EU AI Act, these dimensions group together related provisions into coherent blocks that can later be connected, in a structured way, to metrics, pipelines, controls and documentation artefacts.[1,3] In practical terms, they answer the question: *which parts of the GDPR and the AI Act are most relevant for the TRCM and should systematically give rise to technical evidence in MLOps workflows?*[2,5]

Table 1 summarises the main regulatory dimensions captured by the TRCM and organises them along three complementary questions.[6,20] The first column (*Regulatory dimension*) answers: *What are the core regulatory pillars that matter for high-risk AI systems?* The second column (*Primary legal focus*) specifies: *Which GDPR and EU AI Act provisions concretely anchor each pillar?* The third column (*Illustrative technical evidence*) indicates: *What kinds of technical artefacts and records would constitute convincing evidence that this pillar is being met in practice?*

By framing the table in this way, the regulatory dimensions can be read directly as a bridge between abstract legal duties and the types of evidence that development and governance teams must be able to produce. The remainder of this subsection elaborates how each dimension complements principle-level regulation with concrete, audit-ready development artefacts.[9,11]

 doi:10.20944/preprints202512.0593.v1

Table 1. Regulatory dimensions and their primary legal anchors.

Regulatory dimension	Primary legal focus	Illustrative technical evidence
Data protection and data governance	GDPR data protection principles (lawfulness, purpose limitation, minimisation, accuracy, storage limitation, accountability; Arts. 5–6), special categories and safeguards (Art. 9), information duties and data subject rights (Arts. 13–15), and AI Act data governance requirements for training, validation and testing data (Art. 10).[1,25]	Dataset documentation, lawful-basis and purpose records, data minimisation and pseudonymisation logs, data quality and representativeness reports, access and retention policies.[6,15]
Risk management, robustness, and safety	AI Act obligations to establish and maintain a risk management system and to ensure accuracy, robustness and cybersecurity of high-risk AI systems (Arts. 9 and 15), together with related safety and risk-mitigation duties.[2,20]	Risk registers linked to use cases, scenario-based tests, robustness and stress-testing reports, security assessment and penetration-testing evidence.[10,12,21]
Transparency, explanation, and human oversight	GDPR duties to provide meaningful information about automated decision-making and enable rights of access, explanation and contestation (Arts. 12–15, 22), and AI Act transparency and human-oversight requirements for high-risk AI systems (Arts. 13–14).[3,26]	User-facing notices, explanation logs, explanation quality metrics, records of human review and overrides, decision dossiers documenting human-in-the-loop interventions.[16–18]
Documentation, logging, and record-keeping	GDPR accountability obligations and record-keeping duties (Arts. 5(2), 24, 30, 35–36) and AI Act requirements on technical documentation, logging and conformity assessment files (Arts. 11–12, 18–21).[1,22]	Model and dataset cards, pipeline and configuration records, event logs and logging schemas, conformity assessment documentation, audit trails.[6,24]
Post-market monitoring and incident management	AI Act obligations for post-market monitoring, reporting of serious incidents and malfunctioning, and updating of risk management and mitigation measures (Arts. 61–62 and related provisions).[20,27]	Monitoring strategies and dashboards, alerting rules, incident tickets and reports, post-mortem analyses, change logs and updated risk and conformity documentation.[8,10,13]

The *data protection and data governance* dimension consolidates GDPR principles and AI Act data provisions into a single view of how personal and non-personal data must be handled across the lifecycle of high-risk AI systems.[1,25] In the TRCM, this implies the existence of artefacts such as dataset documentation capturing provenance and purpose, explicit records of lawful bases and purpose limitation, logs of minimisation and pseudonymisation measures, and bias and drift analyses that evidence data quality and representativeness.[6,15] These artefacts allow organisations to demonstrate, during audits or investigations, that data practices are not only described in policies but technically instantiated and monitored.[11,20]

The *risk management, robustness, and safety* dimension reflects AI Act obligations to maintain a documented risk management system and to ensure that high-risk AI systems meet appropriate standards of accuracy, robustness and cybersecurity.[2,20] Within the TRCM, this is translated into expectations for maintained risk registers tied to specific use cases, scenario-based tests that probe

 doi:10.20944/preprints202512.0593.v1

critical failure modes, robustness and stress-testing reports, and security assessments integrated into the development pipeline.[10,12] In other words, risk-related legal duties must materialise as explicit evaluation artefacts, rather than as implicit or ad hoc performance thresholds.[13,21]

The *transparency, explanation, and human oversight* dimension unifies GDPR duties to provide meaningful information about automated decisions and to enable rights of access and contestation with AI Act requirements on transparency and effective human oversight.[3,26] The TRCM operationalises this dimension by requiring user-facing notices that clearly describe the role of AI, systematic logging of explanation outputs, metrics for explanation stability and fidelity, and records of human review, overrides and escalations.[16,17] Decision dossiers that link inputs, model versions, explanations and human actions become central artefacts for evidencing, in practice, that individuals and oversight bodies can understand and challenge decisions.[6,18]

The *documentation, logging, and record-keeping* dimension integrates GDPR accountability obligations with AI Act duties on technical documentation, logging and conformity assessment.[1,22] In the TRCM, this translates into structured, versioned documentation of models, datasets, pipelines and governance decisions, as well as logging schemas and event logs that support ex post reconstruction of system behaviour.[6,24] These requirements are mapped to artefacts such as model and dataset cards, configuration records and audit trails, making accountability demonstrable through reproducible technical evidence rather than solely through narrative policies.[5,23]

Finally, the *post-market monitoring and incident management* dimension captures AI Act requirements that extend beyond deployment, including continuous monitoring of system performance, detection and reporting of serious incidents, and updating of risk mitigation measures.[20,27] The TRCM associates these obligations with monitoring dashboards, defined alert thresholds, incident-reporting workflows, post-mortem analyses and change logs that link updates to models, data or controls back to revised risk and conformity documentation.[8,10] This ensures that ongoing compliance is supported by a living body of technical evidence that evolves with the system and its environment, rather than by static documentation frozen at deployment time.[13,28]

Taken together, these regulatory dimensions provide the legal backbone of the TRCM by specifying which GDPR and AI Act obligations must be backed by concrete technical artefacts and how they should be grouped for operational purposes.[1,6] In the next subsection, the corresponding technical dimensions show how each of these legal pillars is translated into concrete engineering concerns and development artefacts, so that principle-level regulation can be turned into a structured catalogue of audit-ready evidence across the AI lifecycle.[5,8,9]

3.4. Technical Dimensions

The technical dimensions of the TRCM describe how obligations arising from the GDPR and the EU AI Act are translated into concrete engineering concerns and development artefacts across the AI lifecycle.[1,6] Each dimension captures a coherent area in which legal requirements must be operationalised through metrics, pipeline steps, controls and documentation, so that compliance can be demonstrated through verifiable, audit-ready evidence rather than high-level narratives.[5,8] Taken together, these dimensions provide a structured reference for technical teams designing MLOps workflows that are natively aligned with regulatory expectations for high-risk AI systems.[2,9]

Complementing the regulatory view in Table 1, which answers *which blocks of GDPR and AI Act obligations matter for the TRCM and must be demonstrably satisfied*, Table 2 addresses the corresponding technical question: *how are these obligations materialised as concrete technical work and artefacts within MLOps workflows?* In other words, while the regulatory dimensions identify the legal pillars that require evidence, the technical dimensions organise the engineering effort into reusable patterns of data handling, model evaluation, explanation workflows, monitoring, logging and documentation.

Table 2 summarises these core technical dimensions, indicating (i) their primary focus within the development and operation of high-risk AI systems and (ii) the main types of artefacts that should exist under each dimension.[6,20] The subsequent paragraphs detail how these dimensions collectively ensure that regulatory obligations related to data protection, transparency, human oversight,

 doi:10.20944/preprints202512.0593.v1

risk management and auditability are systematically instantiated in development and deployment practices.[3,11]

Table 2. Core technical dimensions in the TRCM.

Technical dimension	Primary focus	Illustrative artefacts
Data governance	Data provenance, quality, representativeness, bias and lifecycle management.[15,25]	Dataset documentation, lineage records, preprocessing logs, bias and drift reports, access and retention policies.[1,6]
Model behaviour and performance	Accuracy, robustness, calibration and evaluation under risk-relevant scenarios.[10,12]	Validation reports, robustness tests, performance dashboards, stress-test scenarios, model versioning records.[11,21]
Transparency, explainability and communication	Provision of meaningful information about automated decisions and support for human understanding and contestability.[16,17]	Explanation configuration files, explanation logs, stability and fidelity metrics, decision dossiers linking inputs, models, explanations and human interventions.[3,18,26]
Monitoring, logging and incident response	Continuous surveillance of system behaviour, detection of anomalies and handling of incidents.[8,10]	Runtime logs, drift and anomaly indicators, alert rules, incident tickets, remediation playbooks, post-mortem reports.[13,20]
Documentation and auditability	Integration of technical evidence into coherent, audit-ready documentation.[6,22]	Model and dataset cards, DPIA and risk assessment summaries, conformity assessment files, governance records, holistic audit trails.[5,23]

The *data governance* dimension ensures that obligations related to lawful processing, purpose limitation, data minimisation and accuracy are instantiated through explicit control over the data lifecycle.[1,25] By requiring artefacts such as dataset documentation, provenance and lineage trails, preprocessing logs and bias and drift analyses, the TRCM makes it possible to evidence how data used in high-risk AI systems satisfy GDPR principles and AI Act data governance requirements in a traceable manner.[6,15]

The *model behaviour and performance* dimension aligns legal expectations of reliability, robustness and risk management with concrete evaluation practices.[10,21] Models must be accompanied by validation reports, robustness and stress tests, and performance dashboards that reflect risk-relevant scenarios and stakeholder requirements, thereby creating a transparent link between regulatory duties and the quantitative evidence used to justify deployment decisions.[11,12]

The *transparency, explainability and communication* dimension addresses duties to provide meaningful information about automated decision-making and to support effective human oversight and contestability.[3,16] Within the TRCM, these duties are operationalised through explanation configurations, systematic logging of explanation outputs and metrics for explanation stability and fidelity, complemented by decision dossiers that bundle inputs, model versions, explanations and human actions for later review.[17,18,26]

The *monitoring, logging and incident response* dimension translates continuous risk management and post-market surveillance obligations into requirements for runtime observability and structured incident handling.[8,20] Required artefacts include comprehensive runtime logs, drift and anomaly indicators, alerting rules and documented incident-response workflows and post-mortem reports,

 doi:10.20944/preprints202512.0593.v1

enabling providers and deployers to show how emerging risks are detected, investigated and mitigated over time.[10,13]

The *documentation and auditability* dimension integrates outputs from all other dimensions into a coherent, audit-ready evidence base that supports conformity assessments, supervisory reviews and internal governance processes.[6,22] Model and dataset cards, DPIA and risk assessment summaries, conformity documentation and governance records are organised so that each GDPR and AI Act obligation mapped in the TRCM can be traced to specific metrics, pipeline steps, controls and logs, thereby complementing regulatory principles with a concrete catalogue of audit-ready development artefacts.[5,8,23]

A cross-cutting feature of these technical dimensions is that they are designed to be instantiated as reusable patterns within MLOps workflows, rather than as one-off compliance checklists.[8,9] By requiring that each mapped GDPR and AI Act obligation be supported by specific metrics, pipeline steps, controls and documentation aligned with the dimensions in Table 2, the TRCM enables organisations to engineer compliance evidence as an intrinsic by-product of development and operation, providing a stable bridge between evolving regulatory expectations and the concrete artefacts produced by technical teams.[5,6]

3.5. Linking Legal and Technical Dimensions

The previous subsections introduced, separately, the regulatory and technical dimensions that structure the TRCM. On the legal side, Table 1 identifies the main GDPR and EU AI Act pillars that must be reflected in high-risk AI systems. On the technical side, Table 2 groups the corresponding engineering work into core dimensions that can be instantiated in MLOps workflows. This subsection makes the connection between these two views explicit, clarifying how each regulatory pillar is supported by one or more technical dimensions and how each technical dimension, in turn, contributes to several legal obligations.

Figure 3 provides a high-level view of these relationships by depicting the TRCM as a bipartite structure linking regulatory and technical dimensions. Nodes on the left correspond to the regulatory pillars in Table 1, nodes on the right to the technical dimensions in Table 2, and the edges represent explicit correspondences showing which technical dimensions generate the evidence required for each regulatory pillar.

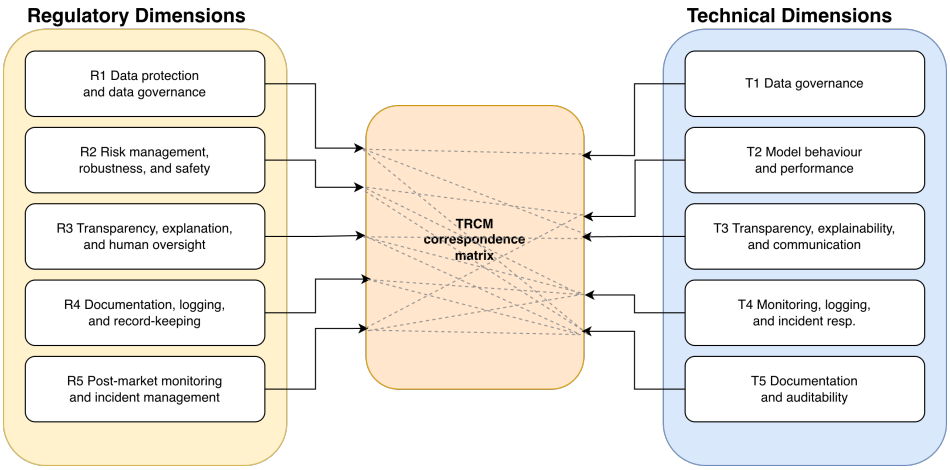

Figure 3. Regulatory–technical links in the TRCM.

Building on this visual overview, Table 3 provides a compact textual summary of the same correspondences. The first column lists the TRCM regulatory dimensions from Table 1; the second column aggregates the technical dimensions from Table 2 that are most directly responsible for generating compliance evidence for each legal pillar; and the third column summarises the rationale for each linkage in terms of evidence production and governance responsibilities. Read from left to

 doi:10.20944/preprints202512.0593.v1

right, the table answers the question: *for this block of GDPR and AI Act obligations, which technical building blocks does the TRCM expect to be present?* Read from right to left, it also clarifies *which legal duties a given technical dimension is primarily designed to support*.

Table 3. Regulatory–technical mapping in the TRCM.

Regulatory dimension	**Main technical dimensions**	**Rationale for the linkage**
Data protection and data governance	Data governance; Documentation and auditability	Data protection principles and AI Act data governance requirements are instantiated through dataset documentation, lineage and preprocessing logs, bias and drift analyses, and retention and access policies, which are then consolidated into audit-ready documentation.
Risk management, robustness, and safety	Model behaviour and performance; Monitoring, logging, and incident response; Documentation and auditability	Risk management and robustness duties require systematic evaluation under risk-relevant scenarios, runtime monitoring for anomalies and failures, and documentation of risk registers, tests and mitigation measures that can be inspected during audits and conformity assessments.
Transparency, explanation, and human oversight	Transparency, explainability, and communication; Monitoring, logging, and incident response; Documentation and auditability	Duties on meaningful information, contestability and human oversight are supported by explanation artefacts and metrics, logging of explanations and human actions, and decision dossiers that preserve a complete record of how decisions were produced, explained and reviewed.
Documentation, logging, and record-keeping	Documentation and auditability; Monitoring, logging, and incident response; Data governance	Accountability, logging and conformity assessment requirements depend on consistent logging across the lifecycle, structured documentation of models, data and pipelines, and the ability to reconstruct how data were collected, processed and used in decision-making.
Post-market monitoring and incident management	Monitoring, logging, and incident response; Documentation and auditability; Model behaviour and performance	Post-market monitoring and incident-reporting obligations rely on continuous observability of system behaviour, incident tickets and post-mortems, and updated documentation that links changes in models, data and controls to revised risk and conformity records.

Taken together, Figure 3 and Table 3 make clear that the TRCM is not a simple one-to-one mapping from legal clauses to isolated technical checks. Rather, it operates as a mesh of correspondences in which each regulatory dimension depends on multiple technical dimensions, and each technical dimension contributes to several legal pillars. By making these linkages explicit, the TRCM helps both legal and technical stakeholders navigate the translation from high-level obligations to concrete engineering work. Legal and compliance teams can start from the regulatory dimensions in Table 1 and, using the mapping, identify which technical dimensions and artefacts need to be in place to

 doi:10.20944/preprints202512.0593.v1

substantiate each pillar. Conversely, engineering and MLOps teams can start from their existing workflows and artefacts and trace, via the same mapping, which GDPR and AI Act duties they are already supporting and where additional controls or documentation are needed. This bidirectional interpretability between legal and technical concerns is what turns the TRCM from a static catalogue of obligations into a practical design and audit instrument for high-risk AI systems.

4. Application of the TRCM to High-Risk AI Systems

Building on the conceptual definition of the Technical–Regulatory Correspondence Matrix (TRCM) in Section 3, this section shows how the matrix can be instantiated and used in practice for concrete high-risk AI systems. Whereas the previous section focused on identifying regulatory and technical dimensions, and on clarifying their many-to-many relationships, the present section takes the perspective of an organisation that must design, deploy and operate a high-risk AI system under the combined obligations of the GDPR and the EU AI Act. The goal is to demonstrate how the TRCM can be used to (i) scope a high-risk use case, (ii) derive the corresponding set of technical controls and evidence artefacts, and (iii) understand the operational implications for different stakeholder groups.

4.1. Selecting a High-Risk Use Case

The application of the TRCM begins with the careful selection and scoping of a use case that clearly falls within the remit of the *high-risk* categories defined in the EU AI Act, and that simultaneously triggers strong obligations under the General Data Protection Regulation (GDPR). Rather than focusing on an isolated deployment, the TRCM is instantiated around a *use-case family* that captures a class of similar systems with shared regulatory and technical characteristics.[20] In this paper, we consider as an illustrative family an AI-based anomaly detection system used by an operator of essential services to monitor network traffic and detect security incidents affecting critical infrastructure. This family combines the operation of critical digital infrastructure with the processing of potentially identifiable network data, and therefore sits at the intersection of AI Act high-risk obligations and GDPR data protection requirements.[6]

The first step is to delineate the system and organisational boundaries of the chosen use case. This includes identifying which assets, processes and decisions are in scope: the sources of data ingested (for example, network flow records, log files, telemetry), the types of automated outputs produced (alerts, risk scores, recommended actions), and the human and organisational roles involved (security analysts, system engineers, data protection officers and external regulators). At this stage, the emphasis is on understanding how the system participates in real decision-making chains: which human actors rely on its outputs, what actions those outputs may trigger, and what harms or benefits could arise from false positives, false negatives or degraded performance.[2]

Once the boundaries are defined, the use case is translated into a *regulatory profile*. On the AI Act side, this profile links the system's characteristics to the relevant high-risk categories in Annex III and identifies the associated obligations (risk management system, data and data governance, technical robustness and accuracy, transparency, human oversight, logging and post-market monitoring). On the GDPR side, the profile highlights the lawful bases for processing, the need for a Data Protection Impact Assessment (DPIA), the applicable data protection principles (lawfulness, fairness, transparency, purpose limitation, data minimisation, accuracy, storage limitation, integrity and confidentiality) and the rights of data subjects that are most salient in the context (in particular, access, rectification, objection and restriction of processing).[8,9]

This regulatory profile acts as the entry point into the TRCM. For each regulatory pillar identified as relevant, the profile signals which cells of the matrix become mandatory, recommended or optional for the given use case. In practice, the outcome of this step is a *filtered matrix*—a subset of the full TRCM that constitutes the minimal set of technical dimensions and evidence artefacts that must be instantiated for the high-risk use case under analysis. This filtered matrix provides a structured, traceable link between the high-level legal obligations and the concrete engineering work that will be

 doi:10.20944/preprints202512.0593.v1

required in the subsequent phases of system design, implementation and operation.[22] A compact summary of the high-risk use-case profile used throughout this section is presented in Table 4.

Table 4. High-risk AI use-case profile for network anomaly detection in critical infrastructures.

Aspect	Description
Application domain	Network monitoring and anomaly detection in an operator of essential services, with potential impact on critical infrastructures and service continuity.
AI Act category (Annex III)	AI systems used for the *management and operation of critical infrastructures*, triggering high-risk obligations (risk management, data governance, robustness, transparency, human oversight, logging and post-market monitoring).
Data processed (GDPR)	Traffic metadata, event logs and other telemetry that can be directly or indirectly linked to natural persons (for example, internal users, administrators, customers).
Main decisions supported	Security alerts, incident prioritisation, triggering of automatic or semi-automatic responses (traffic blocking, host isolation, escalation to incident response teams).
Key risks	False negatives (undetected incidents), massive false positives (analyst overload and alert fatigue), automated actions that may affect rights and freedoms (excessive monitoring, intrusive profiling of internal users).
Salient GDPR obligations	Lawfulness of processing and DPIA, data minimisation and purpose limitation, transparency towards relevant data subjects, integrity and confidentiality of data, management of data subject rights (access, rectification, objection, restriction).
Salient AI Act obligations	Risk management system across the lifecycle, data and data governance, documentation and technical file, robustness and accuracy requirements, transparency and human oversight mechanisms, logging and post-market monitoring, conformity assessment documentation.

4.2. Mapping Technical Controls to Legal Requirements

Once the high-risk use case and its regulatory profile have been established, the next step is to use the TRCM to map concrete technical controls, metrics and artefacts to the corresponding legal requirements. Conceptually, this mapping operationalises the central question posed in Section 3: for each regulatory dimension in Table 1, *which* technical dimensions in Table 2 must be instantiated, and *how* should they be evidenced in the context of the chosen use case? The linking table introduced in Section 3 provides the starting point for this exercise by indicating, at a generic level, which combinations of regulatory and technical dimensions are expected to be relevant.[6]

In the anomaly detection example, the regulatory dimension of *data protection and data governance* is supported primarily by the technical dimensions of *data governance* and *documentation and auditability*. This translates into tangible controls such as data inventories and lineage records for all log sources, formalised retention and deletion policies aligned with GDPR storage limitation, and data quality checks that verify the completeness, accuracy and representativeness of traffic samples used for training and validation. Each of these controls is associated with one or more evidence artefacts (for example, configuration files, data schemas, validation reports, DPIA annexes) that can be linked back to the corresponding cells in the TRCM.[8]

For the regulatory dimensions encompassing *risk management, robustness and safety*, the relevant technical dimensions include *model behaviour and performance*, *monitoring, logging and incident response* and *documentation and auditability*. In practical terms, this means defining and implementing a risk management process that encompasses threat modelling, scenario-based testing, evaluation under realistic network load and noise conditions, and stress-testing for concept drift and performance degradation. The TRCM requires that each risk scenario be linked to specific test suites, performance

 doi:10.20944/preprints202512.0593.v1

thresholds and rollback procedures, and that these links be preserved in the MLOps toolchain so that risk-related evidence can be regenerated or extended whenever the model or its operating environment changes.[5]

The regulatory pillars of *transparency* and *human oversight* map to technical dimensions such as *model behaviour and performance*, *documentation and auditability* and, where applicable, *human–AI interaction design*. In the context of network anomaly detection, this translates into model cards and system cards that document model assumptions, limitations, training data provenance and intended use; explanation artefacts that render alerts interpretable for security analysts (for example, feature attributions, prototypical examples, aggregated explanation dashboards); and operational procedures that define how analysts can contest, override or escalate model outputs. Crucially, the TRCM requires that these explanation and oversight mechanisms be linked back to the relevant GDPR principles (fairness, transparency) and AI Act provisions on human oversight, so that they are not treated as optional user interface enhancements but as compliance-critical design elements.[10]

Table 5 concretises this mapping for the anomaly detection use case, linking each salient regulatory pillar to the corresponding TRCM technical dimensions, example controls and evidence artefacts.

 doi:10.20944/preprints202512.0593.v1

Table 5. Instantiation of the TRCM for the network anomaly detection use-case.

Regulatory pillar	TRCM technical dimension	Illustrative technical control	Evidence artefact
Data protection and data governance (GDPR, AI Act)	Data governance	Inventory of data sources (logs, flows, telemetry) and mapping of personal/pseudonymous data; retention and deletion policies aligned with GDPR storage limitation.	Data asset register, lineage records, retention policies, DPIA annexes.
Data protection and data governance	Documentation and auditability	Systematic recording of datasets used for training, validation and testing; explicit description of sampling and exclusion criteria.	Data sheets, dataset descriptions, quality and representativeness assessment reports.
Risk management, robustness and safety (AI Act)	Model behaviour and performance	Performance tests under different traffic loads, noise levels and drift scenarios; definition of target metrics (for example, minimum detection rate, false positive limits).	Test reports, scenario-by-scenario result grids, documentation of thresholds and acceptance criteria.
Risk management, robustness and safety	Monitoring, logging and incident response	Continuous performance monitoring in production; automatic alerts for degradation beyond predefined limits; documented procedures for model rollback and emergency reconfiguration.	Monitoring dashboards, incident tickets, post-incident review minutes, deployment and rollback logs.
Transparency and human oversight (GDPR, AI Act)	Documentation and auditability	Preparation of model cards and system cards describing goals, assumptions, limitations and acceptable use conditions.	Approved technical datasheets, attached to conformity documentation and operational manuals.
Transparency and human oversight	Model behaviour and performance	Generation of local explanations for alerts (for example, feature contributions, similar past examples) and aggregated views for analysts (explainability dashboards).	Explainability reports, dashboard snapshots, archived examples of explanations linked to incident tickets.
Transparency and human oversight	Human–AI interaction design	Definition of workflows that allow analysts to contest, comment on or override alerts; clear escalation and override rules.	Standard operating procedures, records of human decisions, ticket templates capturing human adjudication.

The mapping exercise is not carried out as a static one-off activity. Instead, it is embedded into the MLOps lifecycle as a series of *correspondence checkpoints*. At each major stage of the pipeline—data acquisition, model training, validation, deployment, monitoring and retirement—the TRCM is used to verify whether the expected technical controls for the applicable regulatory dimensions are present, correctly configured and properly evidenced. Where gaps are detected, they are recorded in the same matrix as outstanding actions, together with the responsible role and target milestone for remediation. This approach allows the TRCM to function both as a design-time checklist and as a living registry of compliance-relevant technical decisions throughout the system's lifecycle.[6,22]

Finally, the mapping process produces a structured *traceability view* that can be exported as part of conformity assessment documentation and audit dossiers. For each regulatory requirement, the view

 doi:10.20944/preprints202512.0593.v1

enumerates the supporting technical dimensions, the concrete controls implemented in the system, the associated evidence artefacts and the points in the lifecycle where they are verified or updated. In the anomaly detection use case, this enables, for instance, a regulator or internal auditor to trace how the obligation to maintain appropriate post-market monitoring and incident reporting is fulfilled via monitoring dashboards, alerting rules, incident tickets, post-incident reports and model update logs, all of which are linked back to specific cells in the TRCM.

4.3. Operational Implications

The systematic application of the TRCM has significant operational implications for organisations developing and operating high-risk AI systems. First, it alters the way in which regulatory compliance is perceived within engineering teams. Instead of being treated as an external constraint checked at the end of the project, compliance becomes an organising principle for the design of data pipelines, model architectures, monitoring strategies and documentation practices. This fosters a shift from ad hoc, document-centric compliance towards an *evidence-by-design* approach, in which the artefacts required for demonstrating conformity are generated as a natural by-product of the development and operation of the system.[8]

Second, the TRCM provides a shared coordination mechanism between legal, risk and technical stakeholders. Because the matrix is expressed in terms of both regulatory and technical dimensions, it facilitates structured dialogues in which data protection officers, compliance teams and engineers can jointly inspect the same artefact and reason about coverage, gaps and trade-offs. For example, a discussion on whether a particular logging strategy is adequate for post-market monitoring obligations can take place directly on the relevant cells of the TRCM, linking specific log fields and retention periods to the underlying AI Act and GDPR provisions. This reduces ambiguity, improves mutual understanding and accelerates the resolution of compliance issues that would otherwise require multiple rounds of translation between legal and technical jargon.[9,22]

Third, by embedding the TRCM into existing MLOps tooling, organisations can improve the scalability and repeatability of their compliance efforts. Once a set of correspondence patterns has been established for a given use-case family (for example, network anomaly detection in critical infrastructures), those patterns can be reused and adapted for new projects with similar characteristics. Template matrices, pre-configured monitoring dashboards, standardised DPIA annexes and reusable explanation reports can all be derived from the TRCM and imported into new pipelines with minimal manual effort. Over time, this leads to a library of organisation-specific TRCM profiles and implementation patterns, which can be refined in response to regulatory updates, audit findings or incident learnings.[6]

Table 6 summarises the main operational implications of TRCM adoption for different stakeholder groups in the organisation.

 doi:10.20944/preprints202512.0593.v1

Table 6. Operational implications of TRCM adoption by stakeholder group.

Stakeholder	Main operational implications of TRCM adoption
Development and MLOps teams	Integration of regulatory obligations as engineering requirements from the outset; need to parameterise pipelines (training, validation, deployment, monitoring) so that evidence artefacts are generated automatically; stronger discipline in versioning models, data and configurations.
Data protection officer (DPO) and compliance teams	Access to a shared technical–regulatory artefact for discussing risks and controls; ability to reuse TRCM profiles and DPIA annex templates across similar projects; improved traceability between GDPR/AI Act obligations and the controls actually implemented in systems.
Risk management and internal audit	More systematic mapping between identified risks, implemented controls and monitoring arrangements; clearer visibility of residual risks and outstanding remediation actions recorded directly in the matrix; easier preparation of internal audit work programmes focused on high-risk cells of the TRCM.
Business owners and product managers	Clearer understanding of the compliance impact of design choices; ability to plan budgets and timelines based on the set of technical dimensions and artefacts required by the filtered TRCM; more informed trade-offs between model complexity, operational flexibility and regulatory burden.
Operators and end-users (for example, security analysts)	More transparent and contestable AI behaviour, supported by explanation artefacts and documented escalation paths; clearer articulation of roles and responsibilities in human–AI decision-making loops; better alignment between tool design and operational practices.
Regulators and external auditors	Availability of a structured, verifiable account of the system's lifecycle and control environment; direct traceability from legal provisions to technical controls and evidence artefacts; more efficient conformity assessments grounded in the TRCM rather than in ad hoc document collections.

However, the operationalisation of the TRCM also entails challenges. It requires an upfront investment in modelling regulatory requirements, aligning them with technical practices and configuring toolchains to capture and retain the necessary evidence. If implemented in a purely bureaucratic manner, the matrix risks becoming a static checklist that adds overhead without delivering real gains in safety or accountability. To avoid this, organisations must treat the TRCM as a *governance infrastructure*: a living artefact that is periodically reviewed, adapted to new guidance and case law, and updated when models, data sources or business processes change.[2,9]

Despite these challenges, the benefits of adopting the TRCM for high-risk AI systems are substantial. It offers a concrete pathway for aligning engineering practice with the complex, evolving obligations of the AI Act and the GDPR, while maintaining a focus on operational effectiveness and risk management. For developers, it clarifies which technical work packages are necessary and how they contribute to compliance; for legal and risk teams, it provides a transparent view of how obligations are operationalised; for regulators and auditors, it offers a structured, verifiable account of the system's lifecycle and control environment. As the following proof-of-concept implementation will illustrate, the TRCM can be integrated into real-world development workflows without disrupting existing practices, and can serve as a foundation for more advanced, tool-supported approaches to compliance-by-design in high-risk AI.[8,20]

5. Discussion

The preceding sections have introduced the Technical–Regulatory Correspondence Matrix (TRCM), characterised its regulatory and technical dimensions, and illustrated its application to

 doi:10.20944/preprints202512.0593.v1

a concrete family of high-risk AI systems under the combined obligations of the GDPR and the EU AI Act. This section discusses the broader implications of this proposal, situating the TRCM in relation to existing work on AI governance and compliance-by-design, examining its potential benefits and limitations for practitioners, and outlining directions for future research and standardisation efforts.

5.1. Positioning the TRCM in the AI Governance Landscape

A growing body of literature has emphasised the need for operational frameworks that bridge the gap between high-level AI governance principles and day-to-day engineering practice.[8,29,30] Existing proposals typically fall into one of three categories. First, there are *principle-driven* frameworks that elaborate ethical or legal requirements (for example, fairness, transparency, accountability) into lists of qualitative criteria for organisations to consider. Second, there are *process-oriented* approaches, which focus on governance structures and oversight mechanisms—such as risk committees, model review boards and bidirectional oversight loops—that should be put in place to ensure responsible AI development and deployment.[9,29] Third, there are *tool-centric* initiatives that concentrate on explainability, model cards, dataset documentation or monitoring techniques as building blocks for compliance and accountability.[6,10]

The TRCM is complementary to these lines of work but occupies a distinct niche. Rather than proposing yet another set of high-level principles, a governance structure or a single technical toolkit, the matrix is designed as a *correspondence layer* that explicitly links regulatory pillars (GDPR principles, AI Act obligations, and—where relevant—requirements from management system standards such as ISO/IEC 42001) to families of technical activities and artefacts in MLOps lifecycles. In this sense, the TRCM can be understood as a concrete instantiation of calls for *observability-driven* AI governance,[30] providing a structured way of deciding which observability mechanisms, documentation practices and evidence artefacts are necessary for a given high-risk use case and how they relate to specific legal obligations.

The case study on network anomaly detection illustrates how the TRCM can coexist with other governance instruments. The matrix does not replace organisational roles (such as the data protection officer or the AI compliance function), existing risk management frameworks, or standard documentation practices. Instead, it provides a common reference artefact through which these actors can coordinate, making explicit which cells of the matrix are considered in-scope, which technical dimensions are expected to support which regulatory pillars, and which evidence artefacts are required to demonstrate conformity. In doing so, the TRCM offers an intermediate level of abstraction that is rich enough to be meaningful for practitioners, yet structured enough to support systematic analysis and automation.

5.2. Implications for Engineering Practice and Compliance

From the perspective of engineering teams, one of the main contributions of the TRCM is to make regulatory requirements *actionable* as technical work packages. By mapping each relevant regulatory pillar to a set of technical dimensions, controls and evidence artefacts, the matrix turns abstract obligations into concrete backlog items: data inventory and lineage instrumentation, robustness and stress-testing suites, monitoring and alerting configurations, explainability dashboards, human–AI interaction design patterns and so forth. The anomaly detection example shows how this mapping can be embedded as correspondence checkpoints in an MLOps pipeline, so that compliance-relevant evidence is produced and updated as a by-product of normal development and operations activities rather than as an afterthought.[6,8]

For compliance and risk teams, the TRCM provides a transparent view of how obligations under the GDPR and the AI Act are operationalised in specific systems. Instead of relying on unstructured document collections or ad hoc interviews with engineers, auditors and regulators can inspect the matrix to see which technical dimensions are claimed to support each obligation, what controls are implemented, what evidence exists and where in the lifecycle it is generated. This can facilitate both internal assurance activities and external conformity assessments, particularly in sectors—such as

 doi:10.20944/preprints202512.0593.v1

critical infrastructure and financial services—where regulators are likely to request detailed technical documentation and evidence.[9,20]

At the same time, the adoption of the TRCM encourages organisations to move towards an *evidence-by-design* posture. Because the matrix is instantiated per high-risk use-case family and integrated with CI/CD and MLOps tooling, it naturally pushes teams to treat logging, documentation, monitoring and explainability not as optional extras but as first-class engineering concerns. The resulting artefacts—data sheets, model cards, incident logs, explanation reports—are not only useful for compliance, but also for improving model quality, diagnosing failures and enabling interdisciplinary collaboration. In this sense, the TRCM aligns the incentives of engineers, compliance professionals and business owners around a shared set of artefacts and processes.

Finally, the TRCM has implications for how organisations plan and prioritise investments in AI governance tooling. By making explicit which technical dimensions and controls are required for different use-case families, the matrix can inform decisions about where to invest in automation (for example, automated data lineage capture, explainability pipelines, centralised evidence repositories) and where manual oversight remains necessary. Over time, as organisations accumulate multiple TRCM instances across projects, they can identify common patterns and gaps in their governance infrastructure, guiding strategic investments in reusable components and shared services.

5.3. Limitations and Threats to Validity

Despite its potential, the TRCM as presented in this paper has several limitations that must be acknowledged. First, the matrix is necessarily a *modelling abstraction* of complex legal and technical realities. The choice of regulatory pillars and technical dimensions, as well as the many-to-many mappings between them, reflects the authors' interpretation of the GDPR, the AI Act and current MLOps practices. Different organisations, sectors or supervisory authorities may prefer alternative decompositions or emphasise additional dimensions (for example, sector-specific safety requirements or organisational culture aspects). While the TRCM is designed to be extensible, this subjectivity implies that it should be treated as a starting point for organisational tailoring rather than as a fixed canonical structure.

Second, the primary proof-of-concept focuses on a specific family of high-risk AI systems in the cybersecurity domain: network anomaly detection for operators of essential services. Although this is a relevant and demanding use case, it does not cover other important high-risk categories under the AI Act, such as credit scoring, recruitment, biometric identification or medical triage. The correspondence patterns identified in this paper may need to be adapted or refined when applied to such domains, particularly with respect to fairness, non-discrimination and domain-specific safety requirements.[2,8] Future empirical studies are required to test the TRCM across a broader range of application areas and organisational contexts.

Third, the integration of the TRCM into existing MLOps toolchains and organisational processes is not trivial. The matrix must be connected to concrete repositories, pipelines and evidence stores if it is to act as more than a static checklist. This entails non-negligible engineering effort, as well as governance decisions about who owns and maintains the matrix over time. In organisations with fragmented tooling or siloed teams, there is a risk that the TRCM remains a paper artefact detached from day-to-day practice.

Fourth, there are potential *threats to validity* in the way the TRCM is evaluated. The proof-of-concept presented in this paper is based on a conceptual instantiation and qualitative reasoning rather than on a longitudinal deployment study. As such, claims about improvements in compliance efficiency, audit readiness or interdisciplinary collaboration remain, at this stage, informed hypotheses rather than empirically validated results. To address this, future work should include case studies in real organisations, measuring not only compliance-related outcomes but also impacts on development velocity, incident response and organisational learning.

 doi:10.20944/preprints202512.0593.v1

5.4. Directions for Future Work

The TRCM opens several avenues for future research and practical development. A first direction concerns the *standardisation* of regulatory pillars and technical dimensions. While this paper proposes an initial set aligned with the GDPR, the AI Act and emerging AI management system standards, further work is needed to harmonise these with sectoral guidelines and supervisory expectations. Collaboration with regulators, standardisation bodies and industry consortia could help refine the matrix and promote its adoption as a reference artefact for high-risk AI systems.

A second direction involves the *automation and tooling* around the TRCM. The matrix is particularly amenable to machine-readable representations: technical dimensions, controls and evidence artefacts can be encoded as metadata in configuration files, CI/CD pipelines and model registries. This suggests the possibility of building TRCM-aware orchestration tools that automatically populate matrix cells based on pipeline events, validate coverage against regulatory profiles, and generate up-to-date traceability views for auditors. Such tools could also support what has been described as observability-driven AI governance,[30] by continuously tracking how changes in models, data or infrastructure affect the coverage of regulatory obligations.

A third avenue relates to the analysis of *regulatory tensions and synergies*, particularly between the GDPR and the AI Act. While this paper focuses on aligning obligations where they clearly intersect (for example, data governance, transparency, human oversight), there are areas where the two instruments may pull in different directions, such as data minimisation versus the need for extensive logging and monitoring for safety and accountability purposes.[20] Extending the TRCM to represent not only correspondences but also tensions—for example, via additional matrix layers or annotations—could support more nuanced design trade-off discussions and inform future regulatory clarifications.

Finally, future work should explore the use of the TRCM beyond individual organisations, for example in *ecosystem-level* governance settings where multiple actors (developers, deployers, service providers, regulators) share responsibilities for high-risk AI systems. In such contexts, the matrix could serve as a shared artefact to negotiate responsibility boundaries, evidence-sharing arrangements and coordinated post-market monitoring activities. Investigating these multi-stakeholder applications would further test the robustness and flexibility of the TRCM concept and contribute to the broader project of building trustworthy, legally compliant AI ecosystems in the European regulatory environment.

6. Conclusion and Future Work

This paper has addressed the practical challenge of aligning the development and operation of high-risk AI systems with the combined obligations of the General Data Protection Regulation (GDPR) and the EU Artificial Intelligence Act (AI Act). While recent years have seen a proliferation of high-level principles, governance frameworks and technical toolkits for responsible AI, organisations still lack concrete mechanisms to translate regulatory requirements into actionable engineering work along MLOps lifecycles.[8,29,30] The Technical–Regulatory Correspondence Matrix (TRCM) proposed in this work is a response to that gap.

The main contributions of the paper can be summarised as follows. First, it introduces the TRCM as a structured correspondence layer that explicitly links regulatory pillars—drawn from the GDPR, the AI Act and emerging AI management system standards—to families of technical dimensions in AI and MLOps pipelines. The matrix captures the inherently many-to-many relationships between legal obligations and technical activities, providing a clear vocabulary for discussing how regulatory requirements are supported by specific engineering practices and artefacts.

Second, the paper instantiates the TRCM for a concrete family of high-risk AI systems: network anomaly detection used by operators of essential services to protect critical infrastructures. By constructing a regulatory profile for this use-case family and deriving a filtered TRCM, the paper shows how high-level obligations on risk management, data governance, robustness, transparency and human oversight can be mapped to tangible controls such as data inventories and lineage, stress-testing

 doi:10.20944/preprints202512.0593.v1

suites, monitoring and incident response procedures, explainability mechanisms and human–AI interaction designs. The case study illustrates how these correspondences can be embedded as checkpoints along an MLOps lifecycle, enabling evidence relevant to GDPR and AI Act compliance to be generated and maintained as a by-product of routine development and operations activities.

Third, the paper analyses the operational implications of adopting the TRCM. It argues that the matrix can serve as a shared artefact for coordination between engineering, compliance, risk and audit functions, supporting a shift from document-centric assurance to an *evidence-by-design* posture. By making explicit which technical dimensions are expected to support each regulatory pillar, and by linking them to concrete evidence artefacts, the TRCM can help organisations plan investments in AI governance tooling, identify gaps in their control environment and build libraries of reusable correspondence patterns for recurring high-risk use-case families.

At a broader level, the TRCM contributes to the emerging field of observability-driven AI governance.[30] It provides a systematic way of deciding which observability mechanisms, documentation practices and monitoring arrangements are necessary for a given high-risk AI deployment, and how they relate to specific legal obligations. For regulators and external auditors, the matrix offers a transparent and verifiable view of how obligations under the GDPR and the AI Act are operationalised in concrete systems; for developers and product owners, it clarifies the technical work packages required to support compliance, enabling more informed trade-offs between model complexity, operational flexibility and regulatory burden.

Future Work. While the TRCM is intentionally designed as a flexible and extensible modelling tool, its current formulation has important limitations. The regulatory pillars and technical dimensions proposed in this paper reflect a particular reading of EU law and current MLOps practice, and the correspondence patterns have been illustrated primarily in the cybersecurity domain, focusing on network anomaly detection. Future work should therefore pursue at least four directions.

First, there is a need for broader empirical validation across diverse high-risk categories beyond cybersecurity, such as credit scoring, recruitment, biometric identification and clinical decision support. Applying the TRCM in these domains would test its robustness, reveal domain-specific requirements (for example, fairness and non-discrimination), and inform refinements to both regulatory and technical dimensions.[2,8] Longitudinal case studies in real organisations would also allow quantitative assessment of the TRCM's impact on compliance efficiency, audit readiness, incident handling and organisational learning.

Second, the matrix should be further aligned with emerging standards and supervisory expectations, including AI management system standards such as ISO/IEC 42001 and sectoral guidelines issued by European and national authorities. Collaborative work with regulators, standardisation bodies and industry consortia could help converge towards shared taxonomies of regulatory pillars, technical dimensions and evidence artefacts, increasing the interoperability and portability of TRCM instances across organisations and ecosystems.

Third, the TRCM is a natural candidate for automation. Because its entries correspond to recognisable artefacts and events in AI lifecycles—datasets, model versions, configuration files, CI/CD steps, monitoring alerts, incident tickets—it lends itself to machine-readable representations and tool support. Future work should explore TRCM-aware orchestration and evidence management tools capable of automatically populating matrix cells from pipeline metadata, checking coverage against regulatory profiles, detecting regressions when models or data sources change, and generating up-to-date traceability views for internal and external stakeholders.

Fourth, the TRCM could be extended to represent not only correspondences but also tensions and trade-offs between regulatory requirements, particularly in areas where the GDPR and the AI Act might pull in different directions. Examples include the balance between data minimisation and extensive logging for safety and auditability, or between strong pseudonymisation and the need for meaningful explanations to affected individuals.[20] Representing such tensions explicitly within

 doi:10.20944/preprints202512.0593.v1

the matrix—for instance through additional layers, annotations or risk scores—could support more nuanced design discussions and feed into future regulatory clarifications.

Finally, future research should investigate the use of the TRCM in multi-stakeholder settings, where responsibilities for high-risk AI systems are distributed across developers, deployers, service providers and regulators. In such contexts, the matrix could serve as a shared artefact for defining responsibility boundaries, evidence-sharing arrangements and coordinated post-market monitoring activities. Advancing these lines of work would not only test and refine the TRCM concept, but also contribute to the wider goal of building trustworthy, legally compliant AI ecosystems in the European regulatory environment.

Author Contributions: Conceptualization, A.G. and A.C.; Methodology, A.G.; Software, A.G.; Validation, A.G. and A.C.; Formal analysis, A.G.; Writing—original draft preparation, A.G.; Writing—review and editing, A.G. and A.C.; Supervision, A.C.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The synthetic dataset and all Python notebooks used to instantiate the framework are provided as supplementary material.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Almeida, F.; Calistru, C. Artificial Intelligence and Its Impact on the Public Sector. *Administrative Sciences* **2021**, *11*, 75. https://doi.org/10.3390/admsci11030075.
2. Taeihagh, A. Governance of artificial intelligence. *Policy and Society* **2021**, *40*, 137–157. https://doi.org/10.1080/14494035.2021.1928377.
3. Rodriguez, J.; Farooq, B. Connecting Algorithmic Transparency and Accountability: A Systematic Review. *AI and Ethics* **2023**, *3*, 567–586. https://doi.org/10.1007/s43681-022-00233-8.
4. Justo-Hanani, R. The politics of Artificial Intelligence regulation and governance reform in the European Union. *Policy Sciences* **2022**. https://doi.org/10.1007/s11077-021-09438-2.
5. Papagiannidis, E.; Mikalef, P.; Conboy, K. Responsible artificial intelligence governance: A review and research framework. *J. Strateg. Inf. Syst.* **2025**.
6. Rodríguez, N.D.; Ser, J.; Coeckelbergh, M.; de Prado, M.L.; Herrera-Viedma, E.; Herrera, F. Connecting the Dots in Trustworthy Artificial Intelligence: From AI Principles, Ethics, and Key Requirements to Responsible AI Systems and Regulation. *Information Fusion* **2023**, *91*, 1–25. https://doi.org/10.1016/j.inffus.2023.05.002.
7. Seger, E. In Defence of Principlism in AI Ethics and Governance. *Philosophy & Technology* **2022**, *35*, 1–8. https://doi.org/10.1007/s13347-022-00544-2.
8. Lu, Q.; Zhu, L.; Xu, X.; Whittle, J.; Zowghi, D.; Jacquet, A. Responsible AI Pattern Catalogue: A Collection of Best Practices for AI Governance and Engineering. *ACM Computing Surveys* **2022**, *55*, 1–38. https://doi.org/10.1145/3533372.
9. Birkstedt, T.; Minkkinen, M.; Tandon, A.; Mäntymäki, M. AI governance: themes, knowledge gaps and future agendas. *Internet Research* **2023**, *33*, 135–170. https://doi.org/10.1108/INTR-03-2022-0172.
10. Siala, H.; Wang, Y. SHIFTing artificial intelligence to be responsible in healthcare: A systematic review. *Social Science & Medicine* **2022**, *301*, 114927. https://doi.org/10.1016/j.socscimed.2022.114927.
11. Nasir, S.; Khan, R.A.; Bai, S. Ethical Framework for Harnessing the Power of AI in Healthcare and Beyond. *IEEE Access* **2023**, *11*, 110073–110093. https://doi.org/10.1109/ACCESS.2023.3313492.
12. Stogiannos, N.; Malik, R.; Kumar, A.; Barnes, A.; Pogose, M.; Harvey, H.; McEntee, M.; Malamateniou, C. Black box no more: a scoping review of AI governance frameworks to guide procurement and adoption of AI in medical imaging and radiotherapy in the UK. *The British Journal of Radiology* **2023**, *96*, 20230744. https://doi.org/10.1259/bjr.20230744.
13. Cheng, L.; Varshney, K.R.; Liu, H. Socially Responsible AI Algorithms: Issues, Purposes, and Challenges. *Journal of Artificial Intelligence Research* **2021**, *70*, 1177–1217. https://doi.org/10.1613/jair.1.12782.

14. Sharma, S. Benefits or Concerns of AI: A Multistakeholder Responsibility. *Futures* **2024**, *154*, 103277. https://doi.org/10.1016/j.futures.2023.103277.
15. Ntoutsi, E.; Fafalios, P.; Gadiraju, U.; Iosifidis, V.; Nejdl, W.; Vidal, M.E.; Ruggieri, S.; Turini, F.; Papadopoulos, S.; Krasanakis, E.; et al. Bias in data-driven artificial intelligence systems—An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **2020**, *10*, e1356. https://doi.org/10.1002/widm.1356.
16. Schwalbe, G.; Finzel, B. A comprehensive taxonomy for explainable artificial intelligence: a systematic survey of surveys on methods and concepts. *Data Mining and Knowledge Discovery* **2021**, *35*, 973–1042. https://doi.org/10.1007/s10618-021-00747-7.
17. Rong, Y.; Leemann, T.; trang Nguyen, T.; Fiedler, L.; Qian, P.; Unhelkar, V.; Seidel, T.; Kasneci, G.; Kasneci, E. Towards Human-Centered Explainable AI: A Survey of User Studies for Model Explanations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2022**. https://doi.org/10.1109/TPAMI.2022.3215790.
18. Abgrall, G.; Holder, A.L.; Dagdia, Z.C.; Zeitouni, K.; Monnet, X. Should AI models be explainable to clinicians? *Critical Care* **2024**. https://doi.org/10.1186/s13054-024-05005-y.
19. Martins, T.; Almeida, A.M.D.; Cardoso, E.; Nunes, L. Explainable Artificial Intelligence (XAI): A Systematic Literature Review on Taxonomies and Applications in Finance. *IEEE Access* **2022**. https://doi.org/10.1109/ACCESS.2022.3151234.
20. Outeda, C.C. The EU's AI act: A framework for collaborative governance. *Internet of Things* **2024**, *25*, 101949. https://doi.org/10.1016/j.iot.2024.101949.
21. Li, B.; Qi, P.; Liu, B.; Di, S.; Liu, J.; Pei, J.; Yi, J.; Zhou, B. Trustworthy AI: From Principles to Practices. *ACM Computing Surveys* **2021**, *54*, 1–38. https://doi.org/10.1145/3457607.
22. Schneider, J.; Abraham, R.; Meske, C.; Brocke, J. Artificial Intelligence Governance For Businesses. *Information Systems Management* **2020**, *37*, 314–329. https://doi.org/10.1080/10580530.2020.1818897.
23. Corrêa, N.; Galvão, C.; Santos, J.; Pino, C.; Pinto, E.P.; Barbosa, C.; Massmann, D.; Mambrini, R.; Galvao, L.; Terem, E. Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance. *Patterns* **2022**, *3*, 100550. https://doi.org/10.1016/j.patter.2022.100550.
24. Kuziemski, M.; Misuraca, G. AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications Policy* **2020**, *44*, 101976. https://doi.org/10.1016/j.telpol.2020.101976.
25. Janssen, M.; Brous, P.; Estevez, E.; Barbosa, L.; Janowski, T. Data governance: Organizing data for trustworthy Artificial Intelligence. *Government Information Quarterly* **2020**, *37*, 101493. https://doi.org/10.1016/j.giq.2020.101493.
26. Lazcoz, G.; Hert, P. Humans in the GDPR and AIA governance of automated and algorithmic systems. Essential pre-requisites against abdicating responsibilities. *Computer Law & Security Review* **2023**, *51*, 105824. https://doi.org/10.1016/j.clsr.2023.105824.
27. Zaidan, E.; Ibrahim, I.A. AI Governance in a Complex and Rapidly Changing Regulatory Landscape: A Global Perspective. *Humanities and Social Sciences Communications* **2024**, *11*, 1–12. https://doi.org/10.1057/s41599-024-02613-2.
28. Bouderhem, R. Shaping the future of AI in healthcare through ethics and governance. *Humanities and Social Sciences Communications* **2024**, *11*, 1–10. https://doi.org/10.1057/s41599-024-02500-w.
29. Morley, J.; Floridi, L.; Kinsey, L.A.; Elhalal, A. From What to How: Guidelines for Responsible AI Governance through a Bidirectional and Iterative Oversight Model. *AI & Society* **2021**, *36*, 715–729. https://doi.org/10.1007/s00146-020-00936-9.
30. Hartmann, D. Observability-Driven AI Governance: A Framework for Compliance and Audit Readiness Under the EU AI Act. *AI & Society* **2024**. https://doi.org/10.1007/s43681-024-00595-3.